

BASEMENT



HD23

17414

0. 1291-82 1982

WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

Employment and Training Programs in the 1970s:
Research Results and Methods

Thomas A. Barocci*

WP 1291-82

April 1982

Revised: August 1982

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

Employment and Training Programs in the 1970s:
Research Results and Methods

Thomas A. Barocci*

WP 1291-82

April 1982

Revised: August 1982

* Associate Professor, Alfred P. Sloan School of Management, M.I.T. Special thanks are offered to William Mates who assisted me in organizing and writing the initial draft of this manuscript. Also, I would like to thank Ernst Stromsdorfer and Terry Hedrick for their comments on an earlier draft; both aided in correcting errors and provided the needed push to be concise. Finally, Carolanne Foilb and Ruth Fuquen who put this in its final form; their contribution far exceeded the chore of typing.

Introduction:

Employment and Training research and evaluation efforts during the 1970's accelerated a trend toward more inclusive and quantitatively sophisticated efforts that began in the 1960's. The consensus developed in the mid-1960's, manifested in the Economic Opportunity Act, regarding an active manpower policy and a belief that the government ought to be involved in correcting labor market imbalances, gave way to a more skeptical attitude. The skepticism emerged from both sides of the political spectrum; conservatives questioning whether the government should be involved and liberals questioning the type of involvement, as program dollars alternately tightened and expanded under several administrations. Often, and increasingly so, foundations for settlements of the disagreements over whether and how much, were expected to emerge from an increasingly sophisticated research establishment. The research was to be utilized by the policy makers to alter, discontinue or extend then current programs.

In some cases the research and evaluation achieved its goals; in others it was ignored. Currently, late 1981, there is a threat, already partially executed, to dismantle the employment and training program network (primarily CETA), which, as far as this author can tell, was not based on a careful look at the results of past research and evaluation, but rather done on a basis of anecdotes and isolated evaluations that happened to conform to the campaign promises and the Reagan administration's political need to balance the budget. The decision may have been correct, but in any case, was not based on a consensus of the millions of dollars worth of research funded during the 1970's. The reason, in part, is that there was and still is no consensus in the research on the effects of the programs, either on individual participants or society as a whole. Whether or not they were a good societal investment has never been answered. Lists of successful programs and participants are easily assembled, as are lists of unsuccessful programs and participants who never found their way out of the poverty cycle. Were the successes and the failures the result of the program design, administrative

shorcomings, inadequate funding, lack of participant motivation and/or the economic environment? All, some or none of the above can be documented by one study or another.

From the point of view of research and evaluation design, huge strides have been made; sometimes at the expense of synthesizing efforts to aide in making good policy. Our methods are increasingly sophisticated, as will be detailed in this chapter, and our results are as disparate as ever. However, a survey allows a forum for generalizations and some will be made here. In spite of the methodological battles that have occurred among the researchers, movement toward a consensus on the efficiency of employment and training programs is building.

The main thrust of research and evaluation efforts in the employment and training area was and still is funded primarily by the Department of Labor's Employment and Training Administration (ETA) and the Department of Health and Human Services. Within the Labor Department, the Office of Research and Development and the Office of Program Evaluation and Research (both within ETA) sponsored the bulk of the research and evaluation efforts. For purposes of simplicity in this review, no differentiation between research and evaluation efforts is offered. The two branches of inquiry relating to employment and training programs address the same issues: program effects on society and individuals, whether the program should be retained or modified and whether its funding level or distribution formulae should be altered.

The investment in evaluation of employment and training programs over the decade of the 1970's is not available. Estimates on a yearly basis have been made by the Office of Management and Budget; for example, they estimated that in 1978 \$140 million was spent on evaluating all social programs. Whatever the overall figure, it is substantial and should have had a significant effect on policy.

The purpose of this review is to first synthesize the methods and results of employment and training (E&T) research and secondly to link that research with the policy issues faced during the 1970's. Finally, the prognosis for the field in the 1980's is presented.

1. Programs Included

E&Ts can be classified in various ways. The scheme used in this paper is meant to facilitate the comparison of different programs. Here, E&Ts are considered to fall into one of the following categories: a). classroom or institutional training (CT), b). on-the-job training (OJT), c). adult work experience (AWE), d). public service employment (PSE), and e). youth programs (YPs).

Both CT and OJT can be characterized as skill training programs, since their avowed purpose is to inculcate specific job-related skills. Both have the genesis of their design in the Manpower Development and Training Act programs, and together they account for the largest single part of the evaluation literature.

In contrast, AWE programs attempt to provide the participant with general work experience rather than specific job skills; the aim is to improve the participant's employability.

PSE programs are a peculiar form of E&T inasmuch as they provide actual jobs to participants, rather than training them for pre-existing jobs -- presumably the "training" in a job creation program will arise from the work itself. Although most PSE programs could be considered simply employment programs, rather than training programs, they are included here because of their presumed training-like effects (e.g., increased long run earnings due to enhancement of skills through experiences on the job.) The major goal of PSE programs is, of course, short run earnings for the participants.

Programs designed expressly for youth will be treated separately. Such programs by definition have a clientele -- and thus possibly an impact -- different from that of other types of programs.

Although this study clearly covers a wide variety of E&Ts, it is impossible in a survey of this scope to include every type of "training" program. By necessity, therefore, formal vocational/technical education programs and apprenticeships are expressly excluded from coverage here.

No mention has been made of CETA as a type of E&T, since the main thrust of CETA has been to decategorize and decentralize the administration of federal E&Ts. Thus, it is an administrative structure rather than a substantive type of program. Studies of the impact of CETA per se are thus studies of the difference that program organization -- not training -- makes. While there have been numerous qualitative studies of the implementation of CETA (see, for example, NAS 1976a and 1976b, Snedeker and Snedeker 1978, and Barocci 1978); there have been few sound quantitative studies of CETA's impact.

B. Program Evaluation

Because this review is concerned with evaluating manpower training programs, we need to define program evaluation. Although this may seem elementary, the fact that E&Ts have been "evaluated" for over 15 years without producing conclusive and confidently usable results suggests that "program evaluation" is not so simple a concept as it sounds. (See for example National Academy of Sciences 1974: 102; Goldstein 1972: 14; and Levitan and Wurzburg, 1979, on the inclusiveness of past studies of E&Ts.)

One can approach this question of defining program evaluation by first stating what it is not. It is not pure research, conducted out of an interest in understanding why programs work as they do (Edwards and Guttentag 1976; Posavac and Carey 1980: 10). While such research is important, it is rarely if ever the

reason why an evaluation is undertaken. Nor is program evaluation the same as program management, where "management" is taken to mean the making of decisions regarding the allocation of program resources. In fact, confusing program evaluation with program management puts the evaluator (often an external consultant) in the position of making value-laden, political decisions that by right are the province of responsible program managers and legislators. (Edwards and Guttentag 1975).

Program evaluation is decision-oriented research, that is, research undertaken to provide the factual support for a forthcoming decision or series of decisions. In Borus' words, "Evaluation is the systematic gathering of information in order to make choices among alternative courses of action." (1979:1). It is a systematic study of the consequences of past decisions; its function is to aid in making future decisions.

One important example of program evaluation's role in decision support is in the funding area. Assuming for the moment that E&Ts are intended to alter income distribution, they may be operated and funded with less than total attention to their monetary payoffs for individuals or for society. It is important that clear benchmarks for expected results be established. If those benchmarks are set by individuals skeptical of program worth, the expectation may be unrealistically high and the political problems of maintaining program funding will be compounded. It is thus important to synthesize past evaluations of E&Ts to set realistic performance expectations. Ideally, these performance standards should include guidelines for program management as well as expected results in terms of increased income and post-program employment continuity. Efficiency and equity considerations must be taken into account. Although various typologies for describing policy evaluation research have been suggested (Borus, 1979; Posavac and Carey, 1980, Katz 1975) the basic distinction seems to be that between process and outcome evaluation. As Cain and Hollister state,

There are two broad types of evaluation. The first, which we call "process evaluation", is mainly administrative monitoring....In sum "process evaluation" addresses the question: Given the existence of the program, is it being run honestly and administered efficiently? A second type of evaluation...may be called "outcome evaluation"... With this type of evaluation, the whole concept of the program is brought into question... (1969:..120-121)

Another distinction should be made here. A program's built-in management information system may provide a great deal of information that is relevant to program evaluation, but such data must always undergo selection and analysis before it will support conclusions about program outcomes or other values of interest. Thus, the focus of this review is on formal program evaluations rather than on the implicit evaluations generated by an on-going information system. In line with the foregoing comments on the different types of PE, it would be desirable for such information systems to routinely capture and report information useful for program evaluation and for formal PEs to suggest concrete ways to improve those information systems. But again, such integration is a goal for future evaluation research rather than a reviewable result of past work.

Section II: Evaluation Methods

1. Individual vs. Social Benefits

Definition of relevant program outcomes is the necessary first step in evaluation. Outcomes can be positive or negative, direct or indirect, immediate or delayed and intended or unintended. The most fruitful method of selection is to emphasize outcomes that indicate progress toward program goals (Mangum and Walsh, 1973: 17). Agreement on program goals is, however, unlikely since goals change as economic and political conditions change. In societal terms Weisbrod (1969: 4) best summarizes this perspective with the goals of allocative efficiency,

distributional equity and economic stability, which is a combined equity-efficiency goal. Others (Perry et al., 1975: 3-4) define the goals in individual terms emphasizing that the programs are designed primarily to increase the earnings ability of participants. Still another view is offered by Hammermesh (1971: 6-7) which combines the societal and individual goals by emphasizing human capital enhancement and the better functioning of labor markets.

Disagreement over goals affects evaluation design. For example, if reducing the unemployment rate were the main goal, the evaluator might simply estimate the impact of the E&T expenditures on the national unemployment rate, while playing little or no attention to the gains accruing to the individual participants. If however, goals centering on human capital enhancement were foremost, no evaluative attention would be put on aggregate labor market questions. The goals do, of course, change over time. MDTA grew out of the area Redevelopment Act, which emphasized retraining or moving people to alleviate structural unemployment and it was quite natural to define a program goal as a reduction in the unemployment rate in the relevant area. With the Great Society emphasis shifted to expanding opportunities for disadvantaged individuals and thus goals were defined more in terms of the labor market experience of disadvantaged individuals. (Perry et al , 1975).

The societal and individual benefit distinction would not be important if societal gains (productivity and output) were exactly measured by individual gains (earnings). This equality does not hold since earnings are only one component of total employee compensation, earnings will not equal the worker's marginal product if manpower shortages exist (Hardin, 1969: 101) and because participants may have to forego other opportunities to participate, which will reduce the net benefits to the individual and to society. Possibly even more important are the losses and gains due to displacement and vacuum effects. A graduate may find a job at the expense of a previously employed worker, thus reducing the net gain to society, while enhancing the graduate's earnings. This is especially true when wages are

rigid on the downside (Kiefer, 1979: 15). Also, a trainee may leave a relatively unskilled position and create a vacancy which may draw a new labor force participant; this "vacuum" effect makes the net gain to society larger than the gains to the individual trainees (Borus et al., 1970: 145). Evaluations that fail to take these effects into account may over- or underestimate the benefits of E & T's (Johnson, 1979).

Over time a consensus has emerged (Ashenfelter, 1978; Borus et al., 1970; Perry et al., 1975) that was incorporated into the Congressional amendments to CETA in 1978, when a phrase was inserted in the Act's goals to state that CETA programs should result in an increase in the earned income of participants. In short, most recent evaluations state that the programs have multiple goals, but have then compressed these into a single measure -- the change in earnings of the participants.

A preferable approach for the future would be to acknowledge that program benefits depend on the perspective of the evaluator and to explicitly construct estimates of benefits from several perspectives. At a minimum the perspectives of the participant, the taxpayer and society should be considered. (See Borus, 1979). For reference purposes, Table 1 is offered; it shows the various benefits which would be relevant from each perspective.

Table 1: Benefits of E&T Programs

<u>Benefit</u>	<u>Perspective</u>		
	<u>Trainee</u>	<u>Taxpayer</u>	<u>Society</u>
I. Output and Wages			
- In-program output	0	0	+
- In-program wages or stipends	+	-	0
- Increased post-program output	0	+	+
- Increased post-program wages	+	-0+	0+
- Increased post-program employment	+	-	0
II. Reduced costs for other social programs			
- transfer payments	-	+	0
- administrative costs	0	+	+
III. Taxes			
- increased tax payments	-	+	0
IV. Psychosocial benefits			
- easing entry into labor force	+	-0+	+
- providing further education	+	-0+	+
- helping the disadvantaged	+	-0+	+
- trainee goodwill	0	+	+
- psychological benefits (self-esteem, sense of security)	+	+	+
- job satisfaction	+	+	+
- improved health	+	+	+
- higher social status	+	0	+
- reduced dependency	+	+	+
- improved family life	+	+	+
V. Socioeconomic benefits			
- reduced unemployment	+	+	+
- increased employment	+	+	+
- more equitable income distribution	+	+0-	+
- increased GNP	0	+	+
- more stable prices	+	+	+
- reduced crime	+	+	+
- increased social stability	+	+	+
- reduced discrimination	+	+	+
- better race relations	+	+	+
- better housing	+	+	+

+ a net benefit

- a net cost

0 neither a net benefit or cost

A second fundamental distinction exists between pecuniary and nonpecuniary benefits, such as job satisfaction and self-esteem. (There are also pecuniary and nonpecuniary social benefits, such as increased aggregate output and increased social stability.) Virtually all discussions of outcome evaluation admit the potential importance of nonpecuniary factors, and virtually all point out the difficulty of evaluating benefits not measured in dollars. There are various reasons for this difficulty. For one thing, some of the most important nonpecuniary benefits are psychological states, e.g., increased job satisfaction and a greater sense of personal worth, which are inherently difficult to measure, requiring the use of psychological tests and subjective interview data. Moreover, while economic benefits can be ranked in terms of their dollar value, it is difficult to rank noneconomic benefits. (Perry et al., 1975: 29; cf. Cain and Hollister 1969: 143.) It is precisely these nonpecuniary benefits that may be the most important program effect for especially disadvantaged groups. (Perry et al., 1975: 34.)

3. Short-run vs. Long-run Benefits

A distinction should be drawn between short-run and long-run benefits; for example, participant earnings three months after program completion and earnings five years later. As Rosen (1975, 1976) and other human capital theorists note, annual earnings fluctuate due to random causes, and what is most desirable is some measure of the individual's lifetime earnings curve. Also, if E&Ts are meant to permanently improve the lot of disadvantaged persons, then the life cycle impact of training on earnings is what the evaluator should estimate. However, what is typically available are data collected shortly after program completion. Evaluation of long-run outcomes raises two conceptually distinct problems: 1) what, if anything, do short-run data tell us about long-run outcomes? and 2) how do we weight outcomes occurring at different points in time?

Only a study that documented a person's entire work life subsequent to program participation could truly be said to provide evidence on the permanence of training-related benefits in relation to that theoretical standard, thus, all of our E&Ts evaluations are short-run. Hence, the problem of inferring long-run outcomes from short-run data is not one that will vanish. Different studies have used different approaches to this problem. One method is to collect data on various "indicator variables" shortly after program termination; these variables, which include number of terminations, number of placements, number employed three months after termination, etc., are routinely reported in the Employment and Training Report of the President, and as Borus notes (1978), these variables are in common use among CETA prime sponsors for program evaluation. But as Borus discovered, almost none of these indicator variables are significantly correlated with long-run earnings, weeks employed, amount of public assistance or unemployment insurance received, or educational attainment.

The clear implication of the research (e.g., Gary and Borus, 1980) is that long-run labor market experience must be measured directly rather than inferred from short-run data, and most of the studies that draw conclusions about long-run trends follow this procedure. However, it should be noted that many of the early studies drew no conclusions about long-run impact, and some of those that did merely assumed (incorrectly) that the initial change in earnings would persist over time (Decision Making Information 1971). On the other hand, Parnes and other users of the National Longitudinal Surveys (NLS) have attempted extensive (10-year) follow-up work on the effects of training and other factors on earnings.

The studies that have done follow-up work have had varying results, ranging from almost total disappearance of training gains to persistence of substantial benefits five years out. The most frequent finding seems to be that although the initial advantage of training erodes over time, some advantage persists as long as five years after training. Prescott and Cooley (1971) found an overall decline in

the earnings advantage of about 10% in the second year after program completion, Farber (1971) found the five-year average advantage to be 60% of the initial impact, and Reid (1976) found the advantage in the fifth year to be about 60% (for males) or 100% (for females) of that in the first year after training. One unanswered question is whether the extent of erosion varies by race or sex: Prescott and Cooley found differences by race, sex, and type of training (CT vs OJT), and Ashenfelter (1978: 56) found erosion of 50% after the fifth year for males but no decline in later years for females. Reid found significant erosion for males, slight erosion for black females, and a widening of the earnings advantage for white females, while Farber found the earnings gains of minorities and women to be at least as long-lasting as those of whites and males. Evidence on the durability of non-pecuniary benefits is non-existent. Follow-on work with the Continuous Longitudinal Manpower Sample (CLMS) will allow for more definite conclusions on the pecuniary benefit durability.

In light of these differing results, we can conclude that inferring long-term benefits from short-run data is unwarranted. We must also conclude that the durability of benefits may vary by race, sex, and possibly type of training. Therefore, the indicated methodology is to directly measure long-run impact in one or more follow-up studies; presumably, at some point the training effect will appear sufficiently persistent or sufficiently attenuated for us to extrapolate to the rest of the individual's work life. Of course, if we perceive the goal of manpower training to be the provision of a "quick fix" for low income individuals, then erosion of the training effect over time would not be of so much concern. Once again, our goals for E&Ts will determine the kind of durability we demand.

The second major problem in assessing the long-term impact of E&Ts is how to weight benefits that occur at different times. The basic notion, central to modern financial theory, is that a dollar of benefits received now is worth more than a dollar received later, both because individuals are present-oriented and because a dollar received now can be invested. The usual procedure for handling this is

to discount benefits received in later years by some pre-determined discount rate, according to the formula $PV(B_t) = B_t / (1+r)^n$, where the left-hand side is the present value of the benefits received at time t, r is the discount rate and n is the length of time. Benefits received in different years can thus be reduced to a common metric and added to permit comparison of different programs.

The determination of the appropriate discount rate is a matter of considerable controversy. A high discount rate greatly reduces the value of benefits received far in the future, while a low rate weights them more heavily. Since program costs (see below) are apt to be incurred "up front," with benefits spread out in time, a high discount rate will result in lower estimates of a program's net present value than a low rate. Many analysts argue for a rate equal to the "market rate" of interest on the grounds that that rate is the best available estimate of how highly society values the future, as well as the best estimate of the opportunity cost of resources withdrawn from private use for public programs. Others maintain that the rates prevailing in market transactions yield underestimates of the social utility of future benefits and accordingly urge the use of a rate lower than the market rate of interest. One technique for accommodating this disagreement is to estimate the program's net present value (or benefit-cost ratio) for a range of discount rates, noting the rate at which the program becomes desirable. The reasonableness of that rate can then be assessed by the decision-maker.

Few studies address these issues rigorously. Borus (1970), Ashenfelter (1978), and others ignore the time value of money and present earnings gains in future years with no indication of how a program with a particular earnings pattern may be compared to alternative uses of society's resources. A few studies offer estimates of present value for a range of discount rates, e.g., Decision Making Information (1971: 7.50). The matter is crucial in a world of limited resources where more programs can be conceived than funded. For example, public housing, national health insurance, manpower training, and enforcement of equal opportunity

laws presumably all benefit disadvantaged members of society -- but to decide which one (or mix) represents the best use of public funds requires some sort of common metric such as a net present value or a benefit-cost ratio. Existing studies of E&Ts are severely deficient in this regard.

4. Costs of E&Ts

The discussion thus far has been couched largely in terms of program benefits. However, the concepts apply equally to program costs. E&Ts like other social programs, have costs that should be taken into account in assessing program performance. Restatement of the entire preceeding discussion in terms of costs is unnecessary; attention will be confined here to a few salient points. Extended discussions of program costs can be found in Borus (1979) and other sources. Like benefits, costs differ according to the perspective being used by the evaluator. The example of stipends received by trainees during training is illustrative; such payments represent a benefit to the trainee, a cost to the taxpayer, and a "wash" (no net benefit or cost) to society as a whole. Table 2 notes the costs that will be relevant from various perspectives.

Table 2: Costs of E&T Programs
Perspective

<u>Cost</u>	<u>Trainee</u>	<u>Taxpayer</u>	<u>Society</u>
I. Foregone opportunities during training			
- market output	0	-	-
- earnings	0-	+	0
- non-market output	-	0	-
- leisure	-	0	-
II. Participation costs			
- travel to training site	-	0	-
- out-of-pocket expenses	-	0	-
III. Program costs			
- instructional salaries	0	-	-
- books and supplies	0	-	-
- physical facilities	0	-	-
- overhead costs	0	-	-
- for OJT: supervisory costs	0	-	-
- for YPs: lost school time	-	-	-
- administrative costs	0	-	-
- costs of program evaluation, audit, etc.	0	-	-
IV. Governmental costs			
- increased transfer payments during training	+	-	0
- central administration	0	-	-
- reduced tax revenues during training	+	-	0
V. Psychosocial costs			
- effort needed for training	-	0	-
- feeling of unfairness in those not selected	-	-	-
- increased competition for skilled work	+	-	0
- decreased competition for unskilled work	0	+	0
- separation from family, friends during training	-	0	-
- racial tension after training at sites and in local communities	-	-	-
- heightened expectations that may not be realistic	-	-	-

- a net program cost
+ a net program benefit
0 neither

Evaluations of E&Ts outcomes rarely; however, attend to costs with the same care devoted to benefits. Many "major" studies do not even include cost data (Decision Making Information, 1971). Others take reported program operating costs as given by the agency and add an estimate of foregone participant earnings, e.g., Hardin and Borus (1966), and Cain and Stromsdorfer (1968 -- overhead costs disregarded). Even the studies that use regression methods to estimate program benefits do not use the same level of analytical sophistication in estimating costs. While this casual approach to cost data is probably good enough for a quick assessment of whether a program "costs more than it's worth," a more thorough study should attempt to gather reliable information on program costs.

5. Relating Costs and Benefits

In theory, the way to relate costs and benefits is quite simple. Either the present value of program costs may be subtracted from the present value of program benefits to yield an estimate of the program's net present value, or the benefits may be divided by the costs to obtain a benefit-cost ratio. The decision rules are equally simple: if the program's net present value is greater than 0 or its benefit-cost ratio is greater than (the two are equivalent), then the present value of the benefits exceeds the present value of the costs, and the program is "worthwhile". If several worthwhile programs are competing for a limited budget, then one undertakes the most worthwhile programs first and continues on the the less worthwhile until the budget is exhausted. (In some cases, program "lumpiness" will neccessitate taking the programs out of order to maximize utility.)

Simple as it may be, "Few aspects of manpower program evaluation have generated as much conflict and cntroversy as cost-benefit analysis" (Perry et al., 1975: 33). The major criticism is that cost-benefit analysis systematically undervalues or totally neglects benefits (and costs) that are difficult to quantify; in general, those will be benefits or costs that are nonpecuniary and

equity focused, rather than pecuniary and efficiency focused. As Somers has put it, "The cost-benefit calculus is only one piece of evidence in the appraisal process, and it may not be the most significant piece of evidence" (cited in Perry et al., 1975: 33).

Most published studies have not undertaken a formal cost-benefit analysis; rather, the usual pattern has been to report the relation between program participation and variables of interest, mainly earnings, wage rates, and LFP. A few do compute a benefit-cost ratio, e.g., Borus et al., (1970) and the studies reviewed in Hardin (1969). The others provide no equivalent summary statistic that would permit comparison of one program with another or of E&Ts as a group with other uses of societal resources. Given the misuses to which quantification can be put, the strictures against cost-benefit analyses have a point; however, the need for measuring and comparing program impacts so that society's resources can be used in the most productive way is also undeniable. It seems preferable to develop decision analysis to the point where social, nonpecuniary and long-term benefits and costs can enjoy the same advantages of quantification that are now most applicable to individual, pecuniary, and short-run factors.

6. Summary

The apparent consensus in the literature can be quickly summarized. The definition of benefits and costs depends on the goals ascribed to particular training programs and on the perspective of the evaluator. While programs have multiple goals, analysts tend to use the change in participant earnings or its components as a proxy for total program benefits, perhaps supplemented by attention to the wage and employment components of earnings. The more recent studies tend to follow up on trainees' later work experience to obtain direct evidence of long-run outcomes, but this follow-up generally does not extend past the fifth year after training. Little attention is paid to nonpecuniary benefits, and societal benefits, as noted, are proxied by individual earnings, which implicitly assumes

that displacement and vacuum effects are unimportant.

Program costs, when considered, are typically derived by adding to direct program operating costs some estimate of foregone trainee earnings. Some studies explicitly relate benefits to costs by computing a ratio, but the practice has been less widespread in recent years, leaving assessments of program worth to the implicit judgement of the decision-maker.

Given that the change in participant earnings is to be taken as a measure of program success, the question arises of how best to study that change. The next section of this chapter describes and evaluates the various research designs that have been employed in past evaluations.

B. Research Design

"Research design" refers to the overall strategy that guides the program evaluation. The evaluation should be designed to provide reliable information about program outcomes, i.e., information useful for decision-makers. This may seem obvious, but apparently it is not. As the National Academy of Sciences noted in 1974:

Manpower training programs have been in existence a little over a decade, yet...little is known about the educational or economic effects of manpower training programs. This is troublesome, especially in light of the fact that about \$180 million have been spent over the past ten years in an attempt to evaluate these programs (1974: 1).

Mangum and Walsh agreed in their 1973 assessment: "After ten years, there is still no definitive evidence one way or the other about MDTA outcomes" (1973: 47).

Writing at about the same time, Goldstein concluded, "Despite substantial expenditure of public funds for research and evaluation, there is only limited reliable information about the impact of training" (1972: 14).

Nor have the intervening years cured this problem, although progress has been made. Writing in 1978, Ashenfelter stated that "it is by now rather widely agreed

that very little is reliably known about the actual effects of these programs" (1978: 47).

The failure of past evaluations to provide reliable information about program outcomes is due in large part to inadequate design of evaluation research. In their classic 1963 work on research design, Campbell and Stanley classified research designs into three categories: non-experimental designs (often called natural experiments), quasi-experimental designs, and true experiments. The three differ greatly on the research strategy employed, the data requirements, and the conclusions that each can support. Of these, the natural experiment is the simplest to initiate but presents the most severe measurement problems. It has been used most widely in outcome evaluations.

1. Natural Experiments

Although there are many variants on the theme, the essence of the natural experiment is the before and after study; the evaluator observes the behavior of the group of interest before and after the occurrence of a "treatment. The basic assumption of this design is that the before observation is the earnings level that would have persisted in the absence of training, so that the difference between post- and pre-training earnings is attributed to the program. For example, the Olympus four-cities study used general pre-program labor market trends as a basis for inferring program impact. (Olympus Research Corporation, 1971, reported by Mangum and Robson, 1973).

This research design was the most widely used in early (pre-CETA) evaluations. Unfortunately, non-experimental designs have a serious flaw: they cannot tell us whether the earnings gains were due to the training program or to some other factor such as overall economic conditions. Another "threat to internal validity" arises from possible biases in the enrollee group: if, for example, enrollees were more talented than non-enrollees to begin with, we would expect their earnings to increase over time. Indeed, given the fact that most people earn

more as they grow older, it is hard to conclude anything about program impact from a non-experimental design such as this. It is probably safe to say that no one today would attempt to evaluate outcomes using what must be regarded as the discredited non-experimental design.

2. True Experiments

The polar opposite from "natural" experiments, in terms of data requirements, difficulty of implementation, and reliability of conclusions, is the true experiment, called by Gilbert et al. the "randomized controlled field trial (1975: 39). That is, out of a sample of persons eligible to enroll, some are assigned at random to participate in the program (the "treatment" group) and some are not (the "control" group). Pre- and post-program earnings, say, are observed and compared. Given this design, and assuming non-contamination of the control group, it is possible to conclude that differences in earnings gains were caused by the training program. The strength of this design springs from the fact that random assignment to treatment or control ensures that there are no systematic differences between the two groups that would bias the results.

There are several reasons why this design has not been more widely used. First, it requires evaluators to "get in on the ground floor"; once the program is under way, it is usually too late to randomly assign persons to participate or not. It is also necessary that the pool of eligibles be larger than the number of available slots; if the number of slots is larger, then it is morally and politically difficult to deny eligibles the right to enter training. The cost of experimentation and a general societal bias against "experimenting on people" are also often cited, although as Gilbert et al. note, true experiments are feasible much more often than we usually think and that, where feasible, they are always to be preferred to non- or quasi-experimental designs (1975). Conlisk further argues that the training context is as well suited to experiment as a social program is likely to be. (1979: 93).

3. Quasi Experiments

On most dimensions of interest -- cost, apparent feasibility, complexity, and conclusiveness of results -- the quasi-experimental design (QED) stands between the natural and the true experiment. Unlike the natural experiment, it uses a comparison group, which increases both its complexity and the strength of its results. Unlike the true experiment, it does not randomly assign individuals to treatment and control, which certainly reduces the conclusiveness of its findings. It has become the generally accepted design for evaluating E&Ts. Because there are numerous ways in which the validity of its results can be compromised, great care must be taken in its execution. There are two main variants of this design: time series (longitudinal) and comparison group (cross-sectional) or a combination of TS and CG-multiple observations on multiple groups.

Even with the time series data we do not know what would have occurred in the absence of training. Earnings gains uncovered in TS studies may actually be due to general societal trends or individual maturation.

Borus and Hamermesh (1978: 136) add that the TS is not useful unless one has a good predictive model of how the process in question unfolds over time. Of particular importance for our purposes is the fact that the NLS does not collect detailed information on the type of training undergone, so that it will not support studies of particular programs. One is driven to the conclusion that while TS studies may serve as a source of hypotheses concerning training effects and may in some cases resolve ambiguities about the direction of causation, they cannot be used as the main research design in an evaluation of program outcomes.

The comparison group design, on the other hand, comes closer to the true experiment and has been the design used in most recent studies. The word "comparison" has been used deliberately instead of "control" to call attention to the fact that individuals are not randomly assigned to treatment or no-treatment status in the QED; instead, statistical techniques are used to control for possible

biases introduced by differences between the non-randomly selected treatment and no-treatment groups. To the degree that these techniques fail to capture all of those differences, the treatment and comparison groups will be similar but not necessarily identical, leaving open the possibility that some factor other than participation in training accounts for observed inter-group differences in, say, earnings. (Technically, the problem is known as the "missing variable" problem -- self selection bias is one type.)

Such bias has been a major problem with evaluations, according to Stromsdorfer (1980: 99). For example, persons applying may be more ambitious than non-applicants; in that case, we would expect their earnings gains to be larger. As Sewell (1969:162) notes, if programs select on the basis of ability, then any change in earnings may simply reflect returns to ability rather than training effects. Perry et al. have also noted the possibility of "creaming" in the selection of trainees (1975: 151). On the other hand, persons accepted might be more disadvantaged than non-acceptees, biasing their earnings gains downwards. Gay and Borus (1980) have noted that participants tend to be younger, less educated, less attached to the labor force, and to have lower pre-training earnings than non-participants.

Because of this problem some students of program evaluation have questioned the entire comparison group design. Campbell and Boruch, for example, call the possibility of selection bias a "fundamental flaw" in the comparison group design (1975: 203); according to them and others, "simple applications of multiple regression, covariance analysis, and matching will usually be inappropriate vehicles for estimating [training] effects" (1975: 209).

There have, however, been various defenses of the comparison group design. Gilbert et al. (1975: 119), while generally critical of non-random comparison groups, state that findings without randomization may be accepted if there is no strong reason to believe that the treatment and comparison groups differ in some

unknown ways and if there is also a strong a priori reason to believe that observed effects are due to the treatment. Cain admits that complete elimination of selection bias is probably impractical (1975: 311-312), but he goes on to suggest a procedure for dealing with it. (1975: 314-315). If the trainee and comparison groups differ in mean pre-training earnings, then Cain's procedure is to apply regression coefficients for the comparison group to the mean values of the trainee group's independent variables to obtain an estimate of the bias.

Finally, it would be noted that recent evaluations combine longitudinal and cross-sectional data to draw on the strengths of each. That is, a series of observations are taken for both trainees and the comparison group, permitting analysis of both gross changes with time series and net changes with comparison groups. This approach in effect unites the work of Parnes and his colleagues with that of Ashenfelter and those who follow his comparison group approach. Although by its nature the quasi-experimental design can never overcome all theoretical objections, it seems likely that this hybrid approach will allow for the strongest possible research design short of a true experiment. As Mangum and Walsh note (1973: 23), the defects of previous studies of E&Ts can only be corrected by evaluations that combine use of a comparison group with long-term longitudinal follow-up.

Seemingly, taking the advice offered by Mangum and Walsh (1973), the U.S. Department of Labor's Office of Program Evaluation and Research initiated a large scale project to collect longitudinal data on CETA enrollees. This data set came to be known as the Continuous Longitudinal Manpower Sample (CLMS) and was begun with the onset of CETA in 1973. The data base collection was (and still is --1981) coordinated by Westat, Inc. (Westat, 1979); it provides a longitudinal description of the participants in all major programs sponsored under CETA. This, of course, still leaves the question of what would have happened to the CETA enrollees in the absence of programs. In order to obtain an estimate of the net impact of the

program on earnings, a comparison group was necessary.

An artificial comparison group was constructed from the Current Population Survey, (coupled with earnings data from the Social Security records), and was used to contrast the experience of CETA program terminees. The matching of the CPS group with the CLMS sample participants involved elaborate disqualification algorithms to best facilitate the equality of background, demographics and labor force experience of the match group with the CLMS sample. (Westat, 1981).

Although it is certainly not without problems, the Westat research using the CLMS represents the most advanced and defensible mechanism for measuring the net impact of the program components on the individual participants. Research regarding the appropriate variables on which to match the CPS and CLMS groups continues, as does work on the elimination of selection bias in the samples. (Director, 1979).

5. A Note on Evaluating PSE Programs

Outcome evaluation is especially difficult for PSE programs. To begin, it is not clear whether such programs are training programs at all, or whether they are better classed as job creation or counter-cyclical revenue-sharing programs. Palmer cites two different goals for PSE programs--combatting cyclical and structural unemployment (1978:7); measures of program impact will differ for the two goals. Fechter (1977: 140) flatly describes PSE programs as job creation rather than training, and it does seem that these programs are aimed more at stimulating aggregate demand than at skill training (aggregate supply) and more at macroeconomic goals than individual training gains.

Second, it is impossible to ignore the issue of substitution/displacement here (Nathan, 1979). At its starkest, assume that a local government agency receives funding for 100 PSE positions but uses all of the money to reduce its own wage and salary expenditures. Then the net number of jobs created may be 0, even though the gross number is 100, and the job creation program becomes revenue

sharing (Jerrett & Barocci, 1979, Fechter, 1977).

Third, the relevance of earnings as a measure of program impact is reduced for PSE, since minimum and maximum wages are fixed by law. For perhaps that reason, most studies of PSE impact have focused on the extent of job creation under PSE: the change in earnings is typically not used as an impact measure. (See, for example, Wachter's comment that measurement of earnings gains is not appropriate in studies of PSE--1979: 286).

Fourth, perhaps because employment rather than earnings has been the focus of PSE evaluations, new methods have been developed to evaluate such programs. For example, Jarrett and Barocci (1979) use Markov chain processes to simulate the probable employment experience of age/race/sex cohorts in the absence of PSE; this can be compared to their actual experience to gauge the extent of displacement and job creation. Perhaps because PSE has emerged as a major policy thrust (and thus a subject of study) only in the mid and late 1970s, there is as yet no consensus on how to evaluate PSE programs: at best, there appears to be agreement that the appropriate methods may differ from those applicable to other E&Ts.

C. Analytic Methods

This section focuses on the analytic methods used in comparison group studies. A well-designed comparison group should allow for direct comparison of the post-training earnings, wage rates, employment stability, etc., of trainees and "controls." There are two reasons why this is not always so. First, it is not practical to match trainees and controls closely enough to eliminate all pre-training differences; most CG studies match on relatively few variables. Secondly, matching may introduce regression artifacts (Campbell and Stanley, 1963: 10-12) and increases the probability of Type II errors (Chen 1971).

These problems mean that one must use more elaborate analytic techniques even with a comparison group design. The two main ways of measuring treatment effects are: analysis of raw or standardized change scores, and analysis of covariance

(multiple regression). As Kenny notes (1975: 346-347), these methods rest on different assumptions and accordingly are appropriate in different circumstances. The basic question is how program participants are selected: if they are chosen solely on the basis of group characteristics, then analysis of change scores is appropriate; but if selection is based on individual as well as group differences, then analysis of covariance is necessary. Where the basis for selection is not known presentation of results from both methods of analysis is indicated. One assumption common to both methods is that the earnings (or wages or employment) function does not change between the measurement of the pre-and post-training scores: if it does, neither of these methods will be successful. If the function is stable with the exception of the error variance, it may be possible to correct for this shift in reliability (Kenny 1975: 355).

Dependent variables. Earnings, wage rates, and labor force participation (hours worked, employment stability, incidence of unemployment) have been the major dependent variables in studies of the effects of training. However, these variables have been measured in different ways: by the absolute post-program level, by the change from the pre-program level, and by the percentage change from the preprogram level. Each appears to have certain advantages. The absolute level of earnings attained after training provides direct evidence of the relevance of training for the reduction of poverty, since earnings levels can be compared to the various official definitions of poverty. Similarly, the amount of earnings gain can be related to the cost per trainee to obtain an estimate of the effectiveness of E&Ts relative to other social programs. The percentage change formulation may be most closely related to participants' own perception of what they have gained from training: a doubling of one's income may be significant quite apart from the absolute numbers involved. Other dependent variables have included occupational attainment, measured by Duncan's occupational scale (Andrisani 1977) or by the median income in the occupation (Freeman 1974).

Each of these measures of training effect has its own problems as well. One general problem is what to use as the base year: use of the year preceeding training tends to overstate training gains. Another general problem, and one not addressed is how to account for inflation, which may push up earnings and wages regardless of training. Suojanen (1977) recognized the problem and noted the theoretical solution--to deflate post-training earnings (or wages) by some appropriate price index. Another problem is the non-homogeneity of earnings gains: it may be easier to move from no earnings at all to, say, \$3000 than from \$2000 to \$5000, since the former involves a change from no labor force participation to some, while the latter indicates a move from a low-skill level to a higher one. (Posavac and Carey 1980: 263-264). The use of the mean change in earnings is also problematic unless we identify the base, since a \$100 gain represents 10 percent over a \$1000 base but only 2.5 percent over a \$4000 base.

It should be pointed out that there is disagreement among researchers in this field as to whether earnings is an appropriate dependent variable. It is argued that the use of annual earnings will confound market transactions with issues of labor/leisure choice and the more transitory effects of unemployment (Griliches, 1977, p. 3). Griliches then argues that the usage of wage rates (per hour or week) are a better measure of returns to schooling; this conclusion however cannot be unilaterally applied when estimating returns to E&Ts since one of their goals relates to lifting people of poverty by increasing the continuity of employment/yearly income.

Functional Forms. Most of the studies of training impact use a generally linear form; the exceptions are easily noted. Human capital variables (such as age, education, and experience) and length of training often have been entered as both quadratic and linear terms to capture possible non-linear effects, e.g., diminishing returns to capital or a U-shaped lifetime earnings curve. Examples include Borus 1978 (age and education squared), Flanagan 1974 (total training

squared), and Borus et al. 1970 (age and education squared).

Dependent variables are sometimes entered as log terms; this permits regression coefficients to be interpreted as percentage effects. An example is Andrisani 1977 (log of earnings and hourly wage). Rosen (1976: 7) cites examples of the use of log terms in human capital studies, but it is not clear that such terms are superior to the ordinary linear terms. Also, a number of studies have used interaction terms such as Sex x Hours Worked (Borus et al. 1970), Education x Hours Worked (*ibid.*), Race x Sex x Training Status (Westat 1979), Education x Percent Women in Occupation (Ferber and Lowry 1976) and the like to capture the combined effects of two factors. This procedure represents an alternative to the estimating of separate equations for different demographic groups.

These few examples do not exhaust the possible functional forms, and it is difficult to discern firm guidelines in the literature on when to depart from linearity. As Ehrenberg notes (1979: 152), the linear form is not sacrosanct; if it is used when the phenomenon is in fact non-linear, the misspecification will introduce bias. Ehrenberg recommends testing for sensitivity to functional form, but while a number of studies estimate alternate equations, few estimate alternate forms of the same equation.

Significance tests. Many of the older E&T studies did not report the significance level for their findings; more recent studies tend to indicate which coefficients are significant at the traditional 1% and 5% levels. It is striking that many published studies have reported R^2 values substantially less than .5, and in some cases less than .1. For example Borus et al. (1970) estimated four earnings equations--their highest R^2 value was .12. Given the rarity of R^2 values higher than .3, it is apparent that the earnings equations being used fail to account for the majority of the variance in "outcomes," which is a major shortcoming in a comparison group study that depends for its validity on the ability to specify (and thereby control for) possible sources of selection or other bias.

An important point that appears again and again is the substantial size of the standard errors of the regression coefficients, including the standard errors on the crucial coefficients of the training variables. For example, Kiefer's 1978 study of MDTA training reports several 95% confidence intervals that include zero, i.e., that standard error was at least half as large as the coefficient. It seems essential to report such information, given the relatively small effects that are typically found. Simply to report, say, a \$400 training gain significant at the 5% level is misleading if the standard error is \$250.

Related to the question of significance is the problem of evaluating statistically significant but numerically small training effects. Suppose, to take an extreme case, that the training gain was \$50, with a 99.5% confidence level and a standard error of 50%. Clearly the training is generating a positive gain--but is a \$50 gain worth all the trouble? Kiefer, in both his 1978 and 1979 studies, argues that training programs should be compared to transfer programs, since the goal of all such programs is the reduction of poverty; presumably, one would compute benefit-cost ratios for training, AFDC, the negative income tax, etc., and choose the program with the highest ratio. (In the case of PSE, one would presumably compare the "training" program to other countercyclical measures; in public works programs, to straight revenue sharing, etc.) A related question is whether we want a small benefit to many people or a large benefit to a few--i.e., how do we assess the variance of training gains? We have as yet no firm answers for these kinds of questions.

A. Enrollee Characteristics

Prior to presentation of measured effects of the programs, it is useful to present a summary of the characteristics of the enrollees in four types of programs--Classroom Training, Adult Work Experience, OJT, and PSE. As shown in Table 3, there have been several notable changes in the characteristics of enrollees during the two years shown (FY76 and FY78).

In general, more disadvantaged clients were enrolled in classroom training and adult work experience than were enrolled in OJT and PSE during both of the years. Most extreme, as expected, was the fact that the PSE positions were given to the most advantaged enrollees, as measured by family income, enrollee income, labor force experience and poverty level definitions. The same pattern holds true in both years covered.

Comparison of the two years reveals that PSE jobs were offered to more disadvantaged enrollees in FY78 than were offered in FY76. The changes in CETA eligibility rules over this period seems to have worked in targetting positions more toward the disadvantaged.

Table 3:

Characteristics of Enrollees in Adult CETA Services
FY 1976/FY 1978 (All figures % of total)

Characteristic	Classroom Training		Adult Work Experience		On-the-job-Training		PSE	
	FY76	FY78	FY76	FY78	FY76	FY78	FY76	FY78
Female	50	60	48	56	35	36	34	38
Age: <21	36	40	10	--	33	36	24	23
22-29	40	36	48	49	40	38	43	42
≥30	24	25	42	51	26	26	34	34
Minority	55	50	40	42	38	32	31	39
H.S. Completer	60	61	64	70	69	69	76	75
Veteran	16	11	20	18	24	21	27	24
Below OMB Poverty Level	66	74	61	77	52	62	44	73
Family Receiving Transfer Benefits	36	35	26	36	20	18	16	26
Family Income <\$6,000	64	57	64	62	54	45	48	56
Enrollee Income <\$1,000	56	50	48	54	43	35	38	44
Predominant Labor Force Status (12 mos. pre-CETA)								
Employed (≥90%)	11	10	15	10	16	17	17	8
Unemployed (≥50%)	37	31	34	39	29	24	27	40
Not in labor force (≥50%)	31	30	27	25	27	25	24	22
Residual	22	29	24	27	28	34	32	31

Source: Westat, Inc., Continuous Longitudinal Manpower Survey.

B. Classroom or Institutional Training (CT)

CT programs have been studied most often; Table 4 displays the results of fifteen studies of the earnings gains of CT trainees, broken down by race and sex. All identified comparison group studies that contained data identified as the earnings gains of CT trainees are included. A number of the studies report results that are not statistically significant, while others contain positive or negative outliers. These facts, in conjunction with the wide range of time covered and the fact that some are national samples and some case studies, preclude inclusion of summary statistics relating to the average gain.

There is obviously great variability in the results, ranging from a loss of \$732 for non-white males in Farber's 1971 study of 1968 trainees, to a gain of \$1,456 for non-white females in Kiefer's 1979 study. It should be noted that the large losses and small gains reported in Farber's studies can be attributed in part to his methods: he used a comparison group drawn from the social security sample, and accordingly, was unable to control for education, a possibly crucial omission. With the exception of Kiefer's work using the OEO/DOL data set, the other studies all report some kind of training gain, generally in the \$300-700 range.

It is probably appropriate to characterize these gains as modest, especially in light of the cost of training. As noted earlier, most evaluations do not devote much care to their cost estimates, typically taking estimates, drawn from program records at face value. For that reason, cost information and benefit-cost ratios are not reported here. However, it seems quite plausible, given the magnitudes of the earnings gains, that costs were at least equal to first-year benefits, which would produce benefit-cost ratios less than one (or, equivalently, net present values less than zero).

Table 4: First Year Earnings Gains from Classroom Training Programs

<u>Study</u>	<u>White Males</u>	<u>White Females</u>	<u>Non-White Males</u>	<u>Non-White Females</u>
Main (1969)	409*	409*	409*	409*
Farber (1971a)	-48	132	129	211
Farber (1979b)	-676	-368	-732	-364
Hardin/Borus (1971)	557	895	1,151	1,095
Prescott/Cookey (1971)	719	527	587	624
Sewell (1971)	0	n.a.	429	n.a.
Smith (1971)	416*	416*	416*	416*
Reid (1976)**	268	571	362	546
Ashenfelter (1978)#	121	443	241	455
Kiefer (1978)+	-619	n.a.	-342	n.a.
Kiefer (1979)++,##	-200	1,316	256	1,456
Goodfellow (1979)++	377@	778	525	737
Cookey et al. (1979)@@	303	318	303	318
Gay/Borus (1980)++	125@	1,132	133@	311
Westat (1980)	400	550	200@	500

* overall estimate

** five-year average

base year is 962; five year average

Kiefer's estimate of quarterly gain x 4

+ Kiefer's estimate of weekly gain x 50

++ race categories are Black and Nonblack

@ not statistically significant

@@ three year average; gains not estimated separately by race

All figures represent positive gains unless otherwise indicated; if not indicated, gain is from base year (usually year before training) to year after training. Only comparison group studies are included in the table.



Five comparison group studies provide information on the longevity of CT benefits. Prescott and Cooley (1971) found that benefits in the second post-training year were 90% of those in the first such year. Farber (1971a) found that the average earnings gain for the first five post-training years was 60% of the first year level. Cooley et al. (1979) calculated three-year averages that were higher than the first year level (146% for men and 163% for women). Reid (1976) and Ashenfelter (1978) also found differences by sex; their findings are compared in Table 5 to those of the other studies.

Table 5: Projected earnings gains

Study	3 yr. avg. as % of yr. 1 level	5 yr. avg. as % of yr. 1 level
Prescott/Cooley(1971)	95% (2 yr. avg.)	n/a
Cooley <u>et al.</u> (1979)	146-163%	n/a
Farber(1971a)	n/a	60%
Reid(1976)	~100%, both sexes	~100% women 70-84% men
Ashenfelter(1978)	~100%, both sexes	~100% women 75% men

Based on these findings, one might expect earnings gains recorded in the first year to persist at least five years out for women but to decay somewhat for men. To take a single example, suppose a median first-year gain for white females of \$485 persisted with no decay for five years; at a discount rate of 7%, the net present value (NPV) of that benefit stream would be \$1,989 (assuming benefits accrue at year's end). Ashenfelter notes that the per-trainee cost for the class of 1964 was about \$1,800, of which a substantial amount represented transfer payments (subsidies to trainees); since transfers have no net social cost, the true cost per trainee is much less than \$1,800, yielding a benefit-cost ratio in excess

of 1.0 and possibly in excess of 2.0. By way of comparison, Hardin (1969: 113) summarizes the results of six early studies that computed benefit-cost ratios: the ratios ranged from 0 to 17.3, with most well in excess of 2. (Hardin used a 10% discount rate and assumed a 10-year benefit stream to standardize the studies.)

The longevity of earnings effects is, in short, difficult to document, even in the most tightly controlled design. The importance of this measurement depends on whether one believes that programs should inculcate permanent additions to the participant's human capital, or simply allow them to attain a job whereby they can begin an earnings stream with a greater slope than would have been the case in the absence of the program. Finally, it is worth noting that the persistence of earnings gains will depend, in great part, on the type of position the trainee attains upon completion and the characteristics of the attained position in terms of advancement and pay opportunities.

It is also worth noting that non-experimental studies typically reported much larger training gains than the comparison group studies. For examples, the Olympus Research Corporation's study in four cities reported earnings gains greater than \$1000 for almost all race/sex/city combinations. Decision Making Information's (DMI's) study of a national sample of MDTA trainees reported gains that were lower but still ranged from \$400 to \$1,400 for trainees who were employed both before and after training. For whatever reasons, use of the comparison group method has resulted in substantially lower estimates of the gains from training; moreover, with the exception of Farber's studies, the later evaluations tend to report lower gains than the earlier studies.

There is relatively little information on which to base a decomposition of earnings gains into wage and employment effects. Most early reports of wage effects have been based on the Employment and Training Administration's operating statistics, which are incomplete and subject to possible bias (exclusion of less successful trainees). Perry et al. (1975) reported that only 15% of the records of

participants from fiscal 1969 to 1972 contained wage data. Of the early comparison group studies that examined wages, Main and Qurin found no wage gains for program graduates (except possibly for women), while Sewell and Smith found wage gains of about 25¢ per hour. However, Sewell's study was limited to a largely black sample in North Carolina, and Smith's was based on extrapolation over time during a period of rapid economic growth and a tight labor market.

More recent work suggests that wage gains comprise a relatively small part of earnings gains. Goodfellow's (1979) study found hourly wage rate changes ranging from a loss of 3¢ per hour for nonblack males, to a gain of 26¢ per hour for nonblack females. However, some of the changes were not statistically significant. Cooley et al. also decomposed earnings gains into wage and employment effects: while their results are not readily translatable into wage rate changes, they suggest that whereas for females the main benefit of training was improved skill levels (reflected in higher wage rates), for males the main benefit was reduced unemployment. While these and the earlier studies do not seem a sufficient basis on which to form a general conclusion, they at least strongly suggest that relatively little of the benefit of CT comes from increased productivity, i.e., higher wages.

The opposite side of the coin is the effect of training on labor force participation (LFP), i.e., the incidence of unemployment and the number of hours worked if employed. Here again the evidence is fragmentary. Main's 1968 study concluded that with no discernible wage effect, the gains from training were almost entirely due to the fact that more of those who completed training were employed; however, it must be pointed out that more completers than controls were high school graduates in the Main study, an obvious source of bias. Non-experimental studies such as the Olympus and DMI studies have also concluded that the largest part of the trainees' earnings gains is due to employment effects. Cooley et al. found that CT increased the probability of employment by four percentage points for men and by

three points for women (1979: 142). Goodfellow (1979) found that CT significantly increased total hours worked and weeks employed and significantly reduced weeks not in the labor force and weeks unemployed (the last for two of the four race/sex groups). Kiefer (1979), however, found numerically large effects only for black female trainees. Thus, while the evidence suggest that training's main effect is to increase LFP (and thereby earnings), the studies are not yet sufficiently broad or conclusive to warrant firm conclusions. The matter is important, however, for if increased trainee employment comes at the expense of other groups in society (i.e., if displacement occurs), then the net gains to society are far less than those suggested by increased earnings alone.

Information on the noneconomic or social impacts of CT is scant. Studies by Main and Gurin tend to show that persons who complete training are more pleased with the program and with their post-program jobs than non-completers, but the use of non-completers (Gurin) and unemployed relatives and neighbors (Main) as comparison groups introduces obvious sources of selection bias. Cohen's 1969 study found that MDTA training programs had a little effect on national unemployment rates, and the Olympus study found little impact on the supply of labor in occupations with a shortage of skilled labor. However, no rigorous comparison group studies of noneconomic or social benefits have been performed.

In short, it seems reasonable to expect earnings gains of \$300-500 from CT, except possibly for white males. These figures ideally should be adjusted to real dollars and then compared in terms of percentage changes. Alterations and sophistication of methodology over the time span covered in this review preclude this type of adjustment; however, the research does not show large "hidden" (non-economic) benefits, and it is unclear whether the benefits of CT exceed the costs, at a reasonable discount rate, especially when the possibility of displacement is taken into account. It is equally unclear whether CT should be assessed solely on the basis of its cost-benefit ratios. There is, however, a

clear need for better cost data and more attention to the displacement- and inflation-related issues before a conclusive assessment of CT is offered.

C. On-the-Job Training (OJT)

Much of what has been said about CT applies also to OJT, as the two types of training are frequently studied together. Table 6 displays the results of past evaluations of OJT, with the same conventions and restrictions as in Table 5. In general, the level of earnings gains is somewhat lower than that for CT, and again white males seem to gain less than other groups. However, Westat's data shows the greatest gains attained, especially for non-whites, attributable to OJT programs. Reports of losses from OJT are limited to Kiefer's 1979 study: one can hypothesize with some certainty that foregone earnings and possibly program costs are smaller for OJT than for CT. There is even less information here on the wage and employment components of earnings gains, on noneconomic and social benefits, and on the duration of earnings gains, leaving those questions open.

D. Adult Work Experience (AWE)

Table 7 summarizes the earnings gains reported in past studies of AWE programs. Interestingly, gains are somewhat higher than those for either of the

Table 6: First Year Earnings Gains from On-the-Job Training Programs

<u>Study</u>	<u>White Males</u>	<u>White Females</u>	<u>Non-White Males</u>	<u>Non-White Females</u>
Farber (1971a)	350	291	551	620
Farber (1971b)	88	104	44	300
Prescott/Cooley (1971)	842	n.a.	755	n.a.
Sewell (1971)	0	756	384	756
Kiefer (1978)+	-619	n.a.	-342	n.a.
Cooley et al. (1979)@@	303	318	303	318
Gay/Borus (1980)++	125@	1,132	133@	311
Westat (1980)	750	550	1,150	1,200

* overall estimate

+ Kiefer's estimate of weekly gain x 50

++ race categories are black and nonblack

@ not statistically significant

@@ three year average; gains not estimated separately by race

All figures represent positive gains unless otherwise indicated; if not indicated, gain is from base year (usually year before training) to year after training. Only comparison group studies are included in the table.

Table 7: Earnings Gains from Adult Work Experience Programs (One Year)**

<u>Study & Program</u>	<u>White Males</u>	<u>White Females</u>	<u>Non-White Males</u>	<u>Non-White Females</u>
Farber (1971c) (JOBS) (contract)	-30	207	188	747
Farber (1971c) (JOBS) (noncontract)	655	623	692	978
Kiefer (1979) (JOBS)	108	1,640	432	728
Goodfellow (1979) (JOBS) (contract)	420 ^o	776	661	463
Gay/Borus (1980) (JOBS)	195 ^o	1,011	646	636
Urban Systems (1971) (CEP)	235*	235*	235*	235*
Leone et al. (1972) (CEP)#	521	187	521	187
Leone et al. (1972) (CEP)##	1,268	-327	1,268	-327
Olympus (1971) (CEP)+	300*	300*	300*	300*
Westate (1980) (CETA)-	450	50	0	300

* overall figure

^o not statistically significant

not a comparison group study; gains are for 1969 participants

not a comparison group study; gains are for 1970 participants

+ not a comparison group study

** Gains are from base year (usually year before training) to year after training.

skill training programs, although several of the studies were non-comparison group efforts. The apparent superiority of this class of programs may be because many AWE programs are in fact "job development" programs; that is, programs that focus mainly on placing enrollees in job openings. For example, the JOBS program provided placement-type services and incentives for private-sector employers to hire unskilled persons, and it is worth noting that some of the largest gains reported in Table 7 are for this program. Moreover, JOBS was created in 1968, a time when the labor market was tight; the results of past studies may thus not generalize to times of labor market slack.

In general, the studies of JOBS cited in Table 7, as well as non-experimental studies not cited, show significant earnings gains for all groups except white males as with other programs for which this result holds, there is no obvious reasons for the failure of that single group to benefit as much as the others. The benefits to women and minorities must be interpreted cautiously: most studies of the program have emphasized the peculiar labor market conditions of the late sixties; some studies have concluded that the jobs obtained by JOBS enrollees could have been gotten without the program (see Comptroller General 1971: 24).

Evidence on the wage and employment components of earnings gains is limited to the studies by Kiefer and Goodfellow. Goodfellow found statistically significant wage gains (27¢ per hour) only for non-black males. On the other hand, Goodfellow found significant effects on total hours worked and weeks employed for three of the four race/sex groups; the group for which the LFP gains were not significant was non-black males. Kiefer found significant effects on labor force participation and the probability of employment only for nonblack females. Goodfellow's results and the general pattern of gains might be explained as follows: if white males had higher pre-program LFP than the other groups and were more able, then they might reap lower employment gains than the other three groups. However, Kiefer's findings do not offer strong support for that hypothesis.

Evidence on the noneconomic impact of JOBS is limited to a study by Greenleigh Associates that surveyed the attitudes of JOBS enrollees. The study found that enrollees attributed more adequate income, greater self-esteem, and improved family life to program participation.

Data on the economic impact of Comprehensive Employment Program (CEP) is scarce, being limited to four studies: Olumpus (1971), Urban Systems (1971, Leone et al. (1972), and Systems Development Corporation (SCD), 1970. The first three studies provide data on earnings gains of CEP enrollees and generally indicate substantial gains for males but not for females. However, neither the Olympus nor the Leone studies used comparison groups. The SDC study gathered information on the wage rates obtained by CEP graduates; again no comparison group was used. SCD's percentile data reveal a general upgrading of wages for both sexes, although the exclusion of those with no work history before or after the program biases the results. It should be noted also that the Urban Systems study was limited to rural CEPs and the Leone study to Philadelphia CEPs. In short, no good evaluation of CEP's economic impact has yet been performed. Evidence on noneconomic impacts is similarly limited and methodologically flawed (see the summary in Perry et al., 1975: 356-359).

D. Youth Programs (YPs)

Two E & Ts that deal specifically with youth are the Job Corps (JC) and the Neighborhood Youth Corps (NYC). Both have been the subject of several comparison group studies, permitting the formation of some tentative conclusions regarding their impact. One problem, however, should be kept in mind: the mere passage of time can convert a teenager who is not in the labor force to one who is, quite apart from participation in any program. It is hard to construct a control group that will avoid this possible maturation bias. (Currently, several studies of the impact of the Youth Incentive Entitlement Pilot Program (YIEPP) under CETA, Title

IV, as amended 1978 are being conducted. For a preliminary baseline report, see Barclay et al., 1979.)

Table 8: Earnings Gains from Youth Programs (One Year)

<u>Study & Program</u>	<u>White Males</u>	<u>White Females</u>	<u>Non-White Males</u>	<u>Non-White Females</u>
Cain (1968) (JC)	203*	203*	203*	203*
Engleman (1971) (JC)++	600	600	535	635
Kiefer (1979) (JC)+,**	-764	-1,012	960	292
Goodfellow (1979) (JC)+	345@	-150@	516@	-71@
Gay/Borus (1980) (JC)+	-215@	-1,282	148@	-325
Borus <u>et al.</u> (1970) (NYC)	136*@	136*@	136*@	136*@
Somers/Stromsdorfer (1970) (NYC)	356#	0@	1,346#	1,084#
Goodfellow (1979) (NYC)+	396@	74@	1,282	56@
Kiefer (1979) (NYC)=,**	-844	-592	1,336	368
Gay/Borus (1980) (NYC)	-896	-1,201	404	1,165

* overall estimate

** Keifer's estimate of quarterly gains x 4

+ race categories are black and nonblack

++ figures for trainees who stayed in the Job Corps 7-12 months; Engleman's estimates were multiplied by 12 to obtain estimates of annual earnings

@ not statistically significant

after-tax gains

before-tax gains (after-tax gains not available)

All figures represent positive gains unless otherwise indicated.

It is also apparent that the various studies have obtained markedly different results; at times even a single study will obtain widely varying earnings gains for the four race/sex groups. Overall, however, the pattern seems to indicate that whites of both sexes experienced small gains or large losses relative to the comparison group, while nonwhites of both sexes achieved small losses or large gains; moreover, men seem to have done generally better than women for both races and both programs.

Part of the explanation of these results may be statistical: the sample sizes in several of the studies were rather small, as evidenced by the unusually high number of statistically insignificant findings (fourteen of the forty entries in Table 8 are not statistically significant). Some light is also shed by decomposition of earnings changes into its components: Kiefer found declines in LFP for the race/sex groups suffering earnings losses, although he could not explain this result. Goodfellow found the same declines in various measures of LFP (total hours worked, weeks worked, weeks not in labor force, and weeks unemployed) for one or more of the race/sex groups; he also noted significant losses of transfer payments for some of the groups.

The suggestion may be that program graduates discover that the post-program employment opportunities available to them are limited and the potential loss of transfer payments large. They may, acting on a rational calculus, drop out of the labor force (temporarily, at least). This would imply some sort of selection bias between trainees and controls, e.g., those entering these programs may be doing so as a last effort to find a worthwhile place in the labor force, while controls may have better employment prospects and thus not view training as necessary. On the other hand, the gains for nonwhite males were substantial, so that training may have beneficial effects for at least that group. Ultimately, however, the fact remains that these studies tend either to find a loss (or a gain not significantly different from zero) or a large gain, suggesting that there may be unexplained

differences in samples, in selection biases, in program content, or in analytic methods that would explain the wide discrepancies in results.

In light of comments that youth programs may be better seen as "riot insurance" than as training programs, information on the noneconomic impacts of these programs would be of particular interest. Unfortunately, the available data consist largely of information on the attitudes of JC and NYC enrollees and graduates; in general, these persons had positive attitudes toward their program and their chances of success in the labor market. (See, for example, Louis Harris and Associates, 1967.) Robin (1969) and others have concluded that youth programs have minimal impact on juvenile delinquency, and Walther et al. (1971) have formed the same conclusion regarding dropping out of high school. It should be mentioned that the YIEPP program (Barclay, 1979) offers a guaranteed job to youth eligibles only if they remain in (return to) school.

A more recent study (Goldberg et al. 1978) is more optimistic: out of 21 measures of noneconomic impact, Job Corpsmen improved on eight relative to dropouts and no shows. The eight were job seeking skills, job satisfaction, attitude towards authority, self esteem, criminal justice system involvement, nutrition behavior, family relations, and leisure time. However, benefits were related to the length of time in the program, and the study's authors noted, "Job Corps must make a concerted effort either to screen out those who seem unlikely to survive the first weeks or to strengthen the program so that more enrollees will remain long enough to benefit." (1978:18) The clear implication is of self-selection bias: completers were evidently different from dropouts and no shows in some unspecified way, leaving open the possibility that they might have improved without the program.

E. Public Service Employment (PSE)

As has already been noted, PSE may require different methods of evaluation than other E & Ts; in any event, many of the studies of PSE have used different methods. As a result, information on the earnings gains of PSE participants is

scarce. There are almost no comparison group studies in this area, the major exception being Westat's 1979 study of PEP. That study compared PEP graduates to a matched sample drawn from the CPS and Social Security Records (Westat, 1979): several different analytic models were employed. As indicated in Table 9, none of the results for white females or the model III results for other males were significant; all other findings were significant at the 1 or 5% levels. Once again, white males showed lower gains or experienced greater earnings declines than the other groups. Westat explained this by noting the superior pre-program labor force experience of white males, e.g., higher earnings and better work history (1979: viii-ix).

Many of the studies of the economic impact of PSE have concentrated on estimating the extent to which the monies provided by the federal government were used as a substitute for local funds, thereby replacing regular employees or paying their salary with federal monies. (The former the displacement effect and the latter the substitution effect.) There is a great deal of disagreement among those who researched this issue on method and meaning. One cannot be sure of the extent of fiscal substitution; however, a review of the studies indicates that the rate is at least 10 percent in the first year and possibly as high as 60 percent. Moreover, it is reasonably certain that the rate of substitution increases with time, as the locality "subsumes" the PSE budget more and more in the regular operating funds, treating it more like general revenue sharing. Finally, it is worth noting that the revised regulations, published in 1977, regarding eligibility and pay levels, have succeeded in reducing the level of substitution. Although there are dozens of studies of this issue, Bassi and Fechter (1979, as revised) summarize them well and, in addition, make their own conclusions and recommendations.

F. Summary

Each category of E & Ts contains very different programs, and comparing

categories is even more of an apples and oranges proposition. However, to give some overall sense of the impact of federal employment and training, Table 11 collects the median one-year earnings gains for the programs; (PSE programs are excluded from this summary). Despite the many problems with the data, the general picture is remarkably stable across program categories and race/sex

Table 9: Earnings Gains of PEP Graduates

<u>Model and</u> <u>Base Year</u>	<u>White</u> <u>Males</u>	<u>White</u> <u>Females</u>	<u>Other</u> <u>Males</u>	<u>Other</u> <u>Females</u>
II; 1969	307	1,127*	552	1,678
II; 1970	317	1,086	704	1,638
III; 1969	-194	373*	144*	870
III; 1970	-181	337*	297*	835

* not statistically significant

Source: Westat, 1979:p. 3-42, Table 3-10

groups. With only one exception (YPs), white males appear to benefit least. For the other groups, AWE seems to be the most beneficial, although the discussion of AWE studies noted possible biases in their results. CT promotes generally larger gains than OJT, and in general YPs are the least beneficial in a strictly economic sense (except for nonwhite males). The range from \$303 to \$576 includes all medians except four, indicating a general first-year benefit level of around \$440. Assuming a five-year life and a 7% discount rate, the present value of that benefit level is about \$1,800. This figure, of course, will vary with the discount rate applied to the calculations.

In closing this review of methods and program impacts, three central points must be reiterated. First, given the limitations in methods, especially regarding the influence of the overall economy, it is uncertain if we can confidently project these results into the 1980s. Persistent inflation, changes in the structure of the labor market, increases in the "natural" rate of unemployment, and the emergence of a "post-industrial" economy are of enough significance to change the operation (and thus expectations) of E & Ts.

Second, assuming that we can satisfactorily project the findings to the next decade, it is still unclear how to evaluate the merit of these programs. What will be the standard of a "successful" program? Is a first-year net earnings increase of \$300-600 enough? Part of the answer must come from the utilization of better cost data in order to compute benefit-cost ratios or net present value. Another part of the answer must come from comparison of the results of E & Ts with other social programs or to the direct and indirect costs of not funding them at all. In the final analysis, the judgment is a political one that can be aided but not determined by the results of research and evaluation.

Finally, assuming that E & Ts pass the economic, equity, and political tests, how can we make them operate better? Estimates of earnings or employment gains for certain demographic groups provide limited help for program managers. In most cases little is known about relating impact data to program content or operations information. We lack "production functions" for programs, primarily because the goals of the programs have never been fully and clearly laid out.

Table 10: Summary Comparison of First Year Earnings Gains for Non-PSE E&T Programs

<u>Program Type</u>	<u>White Males</u>	<u>White Females</u>	<u>Non-White Males</u>	<u>Non-White Females</u>
CT	197	485	333	436
OJT	125	318	303	318
AWE	360	425	536	507
YPS	170	-75	576	248
MEDIAN	184	372	435	377

Outcome evaluation can only tell us whether a program is doing well, compared to some pre-decided benchmarks, but by itself cannot tell us how to improve it. Only persons familiar with the day-to-day operations of a particular program can provide the expertise needed to improve what in the final analysis may be a very important and necessary series of social programs. In making societal decisions regarding programs, legislators, program operators, and designers must consider short and longer run equity side by side with impact and efficiency measures.

Summary: The Implosion of the 1970's:

Writing in 1971, Garth Mangum, emphasized the explosion of manpower research that occurred in the 1960's, bringing it from an "obscure field of interest in 1960 to a major area of academic and commercial effort in 1970" (Mangum, 1971.). In contrast, the research and evaluation done during the 1970's can aptly be described as an implosion. The researchers, inside and outside of academia, took the early work, roundly criticized its methodology, convinced the funding sources of its inadequacy in allowing for certainty in recommendations, and proceeded on a new tack. The new direction was inwardly directed and employed the newest and most quantitatively oriented techniques. Using carefully structured samples, longitudinal followup and statistical analysis packages not previously available, the researchers produced complicated designs and examined the minutiae of resultant data sets.

Research projects aimed exclusively at identification of selection bias in a large data set may be of use to policy makers, but only in the long run and after multiple iterations through academic peers, journal reviewers, public administrators and, finally lawmakers. The latter group most often are inclined to ask only if the results were reasonably accurate. The structure, intended effects and, most importantly, funding levels, for employment and training programs are the result of historical experience, as gauged by carefully structured research efforts

as well as anecdotal accounts of success and failure. Often, the political dimension is most important and overrides other considerations. The decision of the Reagan Administration to cut the funding for CETA programs, particularly PSE, was not the result of careful reviews of past successes of the programs; it was the result of political promises to balance the budget and a lack of cohesive lobbying strategy on the part of program proponents. The entire elimination of heteroscedasticity in earnings models would not have turned the tide.

All of this is not to say that the research of the last decade is of little practical use; rather, the implication is that the focus of research must return to emphasis on the policy implications of the efforts. Lawmakers, federal and state administrators, research firms and academics must all act in concert with full realization that all of the risk associated with investments in public employment and training programs will never be removed. Social programs, especially those that are designed to influence long run earnings and labor market and other experiences, will have participants who do well in the labor market, others who do not. Moreover, the success criteria cannot be narrowly defined by net earnings gain the following year or by placement at completion. Noneconomic and long-run benefits of participation must be considered, as must be the ramifications of elimination of the opportunities offered under the currently operating programs.

Colleagues who have experience working with both public agencies and private firms have been known to remark that they would rather do a cost/benefit analysis for a private employer. When asked why, the response is quite simply, "they are willing to take risks; inordinate time is not spent looking for biases in the data." It would certainly be imprudent to argue in favor of utilizing bad data, but it is equally if not more imprudent to let very good but not excellent data languish as fodder for academic argument rather than offering it and coordinated policy prescriptions to lawmakers and program managers. A major recommendation emerging, after pouring over a seemingly endless series of studies, relates to

where attention of the able research establishment should be put: the assessment of investment risk associated with public training programs. If the programs must pay off in terms of increased earnings (and taxes paid, transfers not paid and the like) this must be explicitly addressed. Parallel to this, the income distributional aspects of employment and training programs must assume a central place in the debate. Even if at a minimum the equity comparison must be with the main alternative, which are transfer programs, the issue should be front and center. As an illustration the prose of Zvi Griliches (1977) and his harking back to more classic literature illustrates the dilemma he found in estimating returns to schooling:

"In a sense, we have circled around our problem and data. We started looking for biases and at first found little. We kept on looking for more and leaned over more until we found ourselves on the other side of the original question. The whole process of such a research venture is perhaps best described by the following conversation between Pooh and Rabbit. (A.A. Milne, The House at Pooh Corner).

"How would it be," said Pooh slowly, "if, as soon as we're out of sight of this Pit, we try to find it again?"

"What's the good of that?" said Rabbit.

"Well," said Pooh, "we keep looking for home and not finding it, so I thought that if we looked for this Pit, we'd be sure not to find it, which would be a Good Thing, because then we might find something that we weren't looking for, which might be just what we were looking for, really."

"I don't see much sense in that," said Rabbit.

"No," said Pooh humbly, "there isn't. But there was going to be when I began it. It's just that something happened to it on the way."

The search for bias in our data on employment training programs should not cease, but rather should be pursued in parallel with our search for better ways to structure and manage employment and training programs. Critiques of the research methods of program evaluation should serve as an input to future evaluation designs, but not as a post-hoc condemnation of outcomes of those efforts. Comments on the limitations of studies should be made concurrent with discussion of their strengths.

The proper utilization of the longitudinal sample of persons participating in CETA is, in my estimation, the analytical issue of most importance to the research and evaluation establishment. Data are still being collected by Westat, Inc., although analysis has slowed, if not ceased. It would be a much regretted decision to cancel this one of a kind data base. Exploration of the data base for use with comparison groups fashioned from the Current Population Survey and the Social Security files is the most fruitful avenue for the econometric research. The problems are well known, but in spite of the yet to be resolved econometric issues, it remains the most comprehensive and usable source of information on the entire employment and training network. There are several ways in which these data can be more effectively utilized: 1. analysis of the longevity of earnings gains, 2. analysis of the usefulness of "placement" rates as an outcome measure, 3. as a tool for analysis of the influence of the economic environment in which the programs operate, as well as their management structure's influence on outcomes.

The CLMS was begun in 1973, coinciding with the commencement of the CETA program. Data are now available for one, two and three year followups of persons who participated in the earlier years. These data can and should be utilized to answer the longer run questions on the returns on investment in the programs. Coincident with this research there is a great need to develop cost data for various programs. Although extremely difficult, estimates are needed if the programs are to be subjected to investment/efficiency criteria, as they often are in political and academic forums.

Placement rates, used as an outcome measure, are subject to a wide variety of reporting and inferential biases. Knowing the percentage of persons leaving any program who are placed in jobs does not tell us about the quality of the positions, nor whether the Prime Sponsor or other placement organization is taking only the

most qualified of those who apply for participation, thereby reflecting the persons' abilities rather than the program effectiveness. These and other reasons, mostly related to the Dept. of Labor's evaluation of Prime Sponsor performance on the basis of placement rates, makes this measure of program effectiveness suspect. Alternatives must be found. One avenue that might be pursued is to establish a clear definition of what constitutes a person's removal from the poverty cycle, possibly including duration of job holding after participation, earnings and degree of reliance on transfer programs. If short term outcome measures such as placement rates, or even wage rates, are to be used, it must be established that the short term indicators have a relationship to long term outcomes that can be categorized and utilized by policy makers.

In some cases it can be strongly argued that the economic environment (and even the social environment) in the geographic area of the program operation can have a significant or even overriding influence on program outcomes. A strong and expanding economy may draw a greater proportion of the CETA participants than a weak economy, regardless of the quality of the programs. Although many studies address this issue, the most valuable source of information on this issue, the CLMS, suppresses geo-specific identification of participants. This suppression was agreed upon at the outset of the data collection (to aid in securing Prime Sponsor cooperation.) This decision should be reconsidered; it is hard to find a justification for Prime Sponsor anonymity in the data base.

Parallel with geo-specific identification, case analyses of the management structure and processes should be undertaken at a sample of Prime Sponsors. Although some ad hoc work on management systems has been undertaken, none has been linked with a reliable sample of participants and their labor force experience. If the data bases would have indicators of management processes as well as location variables, we would be able to respond far more confidently to questions on the determinations of post-program experiences.

The research establishment has come full circle. The debate of the 1960's concerning whether the U.S. should have an active manpower policy was answered in the affirmative. The 1970's brought the research establishment to the technical issues, often at the expense of losing sight of the larger political and theoretical questions. Now, the resources for technical evaluations are drying up quickly; they will not reappear until (if) we reverse policy directions at a later point this decade or this century. In the interim the research establishment will have to devote time to the broader and important theoretical questions, especially those relating to how internal and external labor markets work - a topic left largely unattended over the entire period of active manpower policy in the U.S.

Over the last decade we have learned a great deal about the proper methods of program evaluation and research design. This knowledge must be constructively combined with the unquestionable knowledge that the effects of a social program cannot be predicted with certainty. Dual goals of efficiency and equity preclude complete reliance on returns to investment in employment and training program. The substantial portion of the American public who still live in poverty cannot be expected to be pulled out by an expanding business climate, nor will they be able to lift themselves. Public-private cooperation is the key, and the role of researchers and evaluators must be to facilitate that cooperation. This can be done by designing and carrying through useful research with full knowledge of the political realities of the times and realization that certainty before action might well be solely the preview of the natural sciences.

Bibliography

Adams, Arvil V., "Lessons from the National Longitudinal Surveys: A Commentary," in Current Issues in the Relationship Between Manpower Research and Policy, National Commission for Manpower Policy, Special Report No. 7, 1976.

Andrisani, Paul J., "Internal-External Attitudes, Personal Initiative, and the Labor Market Experience of Black and White Men," Journal of Human Resources, Vol. 12, No. 3, Summer 1977, pp. 308-28.

Ashenfelter, Orley C., "Estimating the Effect of Training Programs on Earnings," The Review of Economics and Statistics, Vol. 60, No. 1, February 1978, pp. 47-57.

Assembly of Behavioral and Social Sciences, National Research Council, National Academy of Sciences, Final Report of the Panel on Manpower Training Evaluation, The Brookings Institution, Washington, DC, 1974.

Baily, Martin N. and James Tobin, "Macroeconomic Effects of Selective Public Employment on Wage Studies," Brookings Papers on Economic Activity, Vol. 1, 1980, pp. 511-539, The Brookings Institution, Washington, DC.

Barclay, Susan, et. al., "Schooling and Work Among Youth from Low Income Households: A Baseline Report from the Entitlement Demonstration," Abt Associates, Cambridge, MA, for the U.S. Department of Labor, Office of Youth Programs, 1979.

Barnow, Burt S., "Theoretical Issues in the Estimation of Production Functions in Manpower Programs," in Bloch, ed., 1979.

Barby, Steven L., Cost-Benefit Analysis and Manpower Programs, (Lexington, MA: D.C. Heath and Company, 1972.)

Bassi, Laurie and Alan Fechter, "The Implications for Fiscal Substitution and Occupational Displacement Under an Expanded CETA Title VI," Technical Analysis Paper 65, U. S. Department of Labor, Office of the Assistant Secretary for Policy, Evaluation and Research, March 1979.

Becker, Gary S., Human Capital, National Bureau of Economic Research, Washington, DC, 1964.

Bennett, C. A. and A. A. Lumsdaine, eds., Evaluation and Experiment: Some Critical Issues in Assessing Social Programs (New York: Academic Press, 1975.)

Blinder, Alan S., "On Dogmatism in Human Capital Theory," Journal of Human Resources, Vol. 11, No. 1, Winter 1976, pp. 8-22.

Blinder, Alan S. and Robert M. Solow, "Analytical Foundations of Fiscal Policy," in The Economics of Public Finance, A.S. Blinder, et. al, eds., The Brookings Institution, Washington, DC, 1974.

3Bloch, Farrell, E., ed., Evaluating Manpower Training Programs (Greenwich, CT: JAI Press, Inc., 1979.)

Boruch, R. F. and H. W. Reichen, eds., Social Experimentation: A Method for Planning and Evaluating Social Intervention (New York: Academic Press, 1979.)

Borus, Michael E., Measuring the Impact of Employment-Related Social Programs: A Primer on the Evaluation of Employment and Training, Vocational Rehabilitation, and Other Job-Oriented Programs (Kalamazoo, MI: The Upjohn Institute for Employment Research, 1979.)

Borus, Michael E., "Indicators of CETA Performance," Industrial and Labor Relations Review, Vol. 32, No 1, 1978, pp. 3-14.

Borus, Michael E., ed., Evaluating the Impact of Manpower Programs, (Lexington, MA: Lexington Books, 1972.)

Borus, Michael E. and Daniel Hamermesh, "Study of the Net Employment Effects of Public Service Employment - Econometric Analysis," in Job Creation Through Public Service Employment, Vol. III, Report #6, National Commission for Manpower Policy, 1978.

Brown, Randall S., Marilyn Moon and Barbara S. Zolath, "Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials," Journal of Human Resources, Vol. 15, No. 1, Winter 1980, pp. 3-28.

Cain, Glen G., "Benefit/Cost Estimates for Job Corps," Institute for Research on Poverty, Madison, WI, 1968.

Cain, Glen G., "Regression and Selected Models to Improve Nonexperimental Comparisons," in Bennett and Lumsdaine, eds., 1975.

Cain, Glen G. and Robinson G. Hollister, "Evaluating Manpower Programs for the Disadvantaged," in Somers and Wood, eds., 1969.

Cain, Glen G., "The Challenge of Segmented Labor Market Theories to Orthodox Theory: A Survey," Journal of Economic Literature, Vol. 14, 1976, pp. 1215-57.

Campbell, Donald P. and R. F. Boruch, "Making the Case for Randomized Assignment to Treatments by Considering the Alternatives," in Bennett and Lumsdaine, eds., 1975.

Campbell, Donald P. and R. F. Boruch, "Reforms as Experiments," American Psychologist, Vol. 24, 1969, pp. 409-29.

Campbell, Donald P. and J. D. Stanley, "Experimental and Quasi-Experimental Designs for Research on Training," in N. L. Gage, ed., Handbook of Research on Teaching, N. L. Gage, ed. (Chicago: Rand McNally, 1963.)

Carliner, Geoffrey, "Returns to Education for Blacks, Anglos, and Five Spanish Groups," Journal of Human Resources, Vol. 11, No. 2, Spring 1976, pp. 172-184.

Cohen, Malcolm, "The Direct Effects of Federal Manpower Programs in Reducing Unemployment," Journal of Human Resources, Vol. 4, No. 4, Fall 1969.

Caro, F., "Evaluation Research: An Overview," in Readings in Evaluation Research, F. Caro, ed. (New York: Russell Sage Foundation, 1971.)

Chen, M. K., "Critical Look at the Matching Technique in Experimentation," The Journal of Experimental Education, Vol. 35, 1967, pp. 95-98.

Chow, Gregory G., "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," Econometrics, Vol. 28, 1960, pp. 591-605.

Comprehensive Employment and Training Act of 1973, Statutes at Large, Vol. 87, pp. 839-83, U.S. Code, Title 29 (Sec. 801, et. al.)

Comprehensive Employment and Training Act Amendments of 1978, Statutes at Large, Vol. 92, pp. 1909-2021, U.S. Code, Title 29 (Sec. 801, et. al.)

Conlisk, John, "Choice of Sample Size in Evaluating Manpower Programs," in Bloch, ed., 1979.

Congressional Budget Office, Temporary Measures to Stimulate Employment: An Evaluation of Some Alternatives, U.S. Congressional Budget Office, Washington, DC, September 1975.

Cook, T.D. and Donald T. Campbell, "The Design and Conduct of Quasi-Experiment and True Experiments in Field Settings," in Handbook of Industrial and Organizational Psychology, M. D. Dunnette, ed. (Chicago: Rand-McNally, 1976.)

Cooley, Thomas F., T. S. McQuire and Edward G. Prescott, The Impact of Manpower Training on Earnings: An Econometric Analysis, U. S. Department of Labor, Employment and Training Administration, Office of Program Evaluation and Research, Washington, DC, September 30, 1975.

Decision Making Information, MDTA Outcomes Study: Final Report, C.A.L. 778, U. S. Department of Labor, Manpower Administration, Office of Program Evaluation and Research, Washington, DC, 1971.

Director, Steven M., "Underadjustment Bias in the Evaluation of Manpower Training," Evaluation Quarterly, Vol. 3, No. 2, May 1979, pp. 190-218.

Duncan, Greg J., "Earnings Functions and Nonpecuniary Benefits," Journal of Human Resources, Vol. 11, No. 4, Fall 1976, pp. 462-83.

Eckaus, Richard, Estimating the Return to Education: A Disaggregated Approach, Carnegie Commission on Higher Education, 1973.

Edwards, W. and M. Guttentag, "Experiments and Evaluations: A Reexamination," in Bennett and Lumsdaine, eds., 1975.

Engleman, Stephan R., "An Economic Analysis of the Job Corps," Ph.D. Dissertation, University of California, Berkeley, CA, 1971.

Farber, David J., "Highlights: Some Findings from a Follow-up Study of Pre- and Post-Training Earnings History of 215,000 Trainees Participating in Two 1964 and Four 1968 Training Programs," U. S. Department of Labor, Manpower Administration, Office of Program Evaluation and Research, Washington, DC, 1971.

Farber David J., "An Analysis of Changes in Earnings of Participants in Manpower Training Programs," U. S. Department of Labor, Manpower Administration, Office of Program Evaluation and Research, Washington, DC, 1971 (cited as 1971b.)

Fechter, Alan, "Job Creation Through Public Service Employment Program," in Job Creation: What Works?, Robert Taggart, ed. (Salt Lake City, Utah: Olympus Publishing Co., 1977.)

Ferber, Marianne A. and Helen M. Lowry, "The Sex Differential in Earnings: A Reappraisal," Industrial and Labor Relations Review, Vol. 29, No. 3, 1976, pp. 377-87.

Finifter, David H., An Analysis of Two Year Post-Program Earning Paths of CETA Participants Using the Early CLMS Cohorts (January 1975-June 1975 Entry), U. S. Department of Labor, Employment and Training Administration, Office of Program Evaluation and Research, Washington, DC, December 1980.

Flanagan, Robert J., "Labor Force Experience, Job Turnover, and Racial Wage Differentials," Review of Economics and Statistics, Vol. 55, November 1974, pp. 521-29.

Freeman, Richard B., "Occupational Training in Proprietary Schools and Technical Institutes," Review of Economics and Statistics, Vol. 56, August 1974, pp. 310-18.

Friedman, Milton, "A Monetary Theory of Nominal Income," Journal of Political Economy, Vol. 79, No. 2, March-April 1971, pp. 323-37.

Garfinkel, Irwin and John L. Palmer, "Issues, Evidence, and Implications," in Creating Jobs: Public Employment Programs and Wage Subsidies, J. L. Palmer, ed., The Brookings Institution, Washington, DC, 1978, pp. 1-42.

Gay, Robert S. and Michael E. Borus, "Validating Performance Indicators for Employment and Training Programs," The Journal of Human Resources, Vol. 15, No. 1., Winter 1980, pp. 29-48.

Geraci, Vincent J. and C. T. King, "Employment and Training (CETA) Program Performance: Long-Term Earning Effects and Short-Term Indicators," paper presented at the Allied Social Science Association Meetings, Denver, CO, September 1980.

Gilbert, J. P., et. al., "Assessing Social Innovations: An Empirical Base for Policy," in Bennett and Lumsdaine, eds., 1975.

Ginzberg, Eli, ed., Employing the Unemployed (New York: Basic Books, 1980), pp. 1-24.

Goldberg, J. F., The Noneconomic Impact of the Job Corps, R&D Monograph 64, prepared by Abt Associates for the U. S. Department of Labor, Employment and Training Administration, Program for Evaluation and Research, Washington, DC, 1978.

Goldstein, J. H., The Effectiveness of Manpower Training Programs: A Review of Research on the Impact on the Poor (Washington, DC: Government Printing Office, 1972.)

Goodfellow, G. P., "Estimates of Benefits of Training for Four Manpower Training Programs," in Bloch, ed., 1979.

Gottman, J. M. and G. V. Glass, "Analysis of Interrupted Time-Series Experiments," in Single Subject Research: Strategies for Evaluating Change, T. R. Kratochwill, ed. (New York: Academic Press, 1978.)

Gramlich, Edward M., "Stimulating the Macro Economy Through State and Local Governments," American Economic Review, Vol. 69, No. 2, May 1979, pp. 180-85.

Greenleigh Associates, Inc., "The Job Opportunities in the Business Sector Program - Are Evaluations of Impact in Ten SMSA's," U.S. Department of Labor, Manpower Administration, Program for Evaluation and Research, Washington, DC, 1970.

Griliches, Zvi, "Estimating Returns to Schooling: Some Econometric Problems," Econometrica, Vol. 45, 1977, pp. 1-22.

Gurin, Gerald, A National Attitude Survey of Trainees in MDTA Institutional Programs, University of Michigan, Ann Arbor, MI, 1970.

Gwartney, Jarvis D. and James E. Long, "The Relative Earnings of Blacks and Other Minorities," Industrial and Labor Relations Review, Vol. 21, No. 3, April 1978, pp. 336-46.

Hamermesh, Daniel S., Economic Aspects of Manpower Training Programs (Lexington, MA: D. C. Heath and Company, 1971.)

Hardin, Einar, "Benefit-Cost Analyses of Occupational Training Programs: A Comparison of Recent Studies," in Somers and Wood, eds., 1969.

Hardin, Einar and Michael Borus, The Economic Benefits and the Costs of Retraining (Lexington, MA: D. C. Heath and Company, 1971.)

Hargrove, E. C., "The Bureaucratic Politics of Evaluation: A Case Study of the Department of Labor," Public Administration Review, March/April 1980, pp. 150-59.

Harrison, Bennett, Education, Training and the Urban Ghetto, (Baltimore, MD: John Hopkins University Press, 1972.)

Haworth, Joan G., James Gwartney and Charles Haworth, "Earnings, Productivity, and Changes in Employment Discrimination During the 1960's," American Economic Review, Vol. 65, No. 1, March 1975, pp. 158-68.

Heckman, James, "Simultaneous Equations Models with Both Continuous and Discrete Endogenous Variables with and without Structural Shifts in the Equations," unpublished, August 1975.

Holton, E. T. and A. A. Lumsdaine, "Field Trial Designs in Gauging the Impact of Fertility Planning Programs," in Bennett and Lumsdaine, eds., 1975.

Jerrett, Robert and Thomas A. Barocci, Public Works, Government Spending and Job Creation (New York: Praeger Press, 1979.)

Johnson, George E., "The Labor Market Displacement Effect in the Analysis of the Net Effect of Manpower Training Programs," in Bloch, ed., 1979.

Johnson, George E. and James D. Tomola, "The Fiscal Substitution Effect of Alternative Approaches to Public Service Employment Policy," Journal of Human Resources, Vol. 12, No. 1, Winter 1977, pp. 3-26.

Kalacheck, Edward and Fredric Raines, "The Structure of Wage Differences Among Mature Male Workers," Journal of Human Resources, Vol. 11, No. 4, Fall 1976, pp. 484-506.

Katz, D., "Feedback in Social Systems: Operational and Systematic Research on Production, Maintenance, Control and Adaptive Functions," in Bennett and Lumsdaine, eds., 1975.

Kemper, Peter and Philip Moss, "Economic Efficiency of Public Employment Programs," in Creating Jobs: Public Programs and the Wage Subsidies, J. L. Palmer, ed., The Brookings Institution, Washington, DC, 1978.

Kenny, D. A., "Quasi-Experimental Approach to Assessing Treatment Effects in the Nonequivalent Control Group Design," Psychological Bulletin, Vol. 82, No. 3, 1975, pp. 345-62.

Kesselman, Jonathan, "Work Relief in the Great Depression," in Creating Jobs: Public Programs and Wage Subsidies, J. L. Palmer, ed., The Brookings Institution, Washington, DC, 1978.

Kiefer, Nicholas M., The Economic Benefits from Four Employment and Training Programs (New York: Garl and Publishers, 1979.)

Keifer, Nicholas M., "Federally Subsidized Occupational Training and the Employment and Earnings of Male Trainees," Journal of Econometrics, Vol. 8, 1978, pp. 111-25.

Killingsworth, Charles C., "The Role of Public Service Employment," in Industrial Relations Research Association: Proceedings of the 1977 Annual Spring Meeting, J. L. Stern and B. C. Dennis, eds. (Chicago: Commerce Clearing House, 1977.)

- King, Alan G. and Knapp, Charles B., "Race and the Determinants of Lifetime Earnings," Industrial and Labor Relations Review, Vol. 31, No. 3, April 1978, pp. 347-55.
- King, Randall H., "Some Further Evidence on the Rate of Return to Schooling and the Business Cycle," Journal of Human Resources, Vol. 15, No. 2, Spring 1980, pp. 264-72.
- Kirk, R. D., Experimental Design: Procedures for the Behavioral Sciences (Belmont, CA: Brooks/Cole, 1968.)
- Kohen, Andrew I., "Determinates of Labor Market Success Among Young Men: Face, Ability, Quantity and Quality of Schooling," Center for Human Resource Research, Ohio State University, Columbus, OH, 1973.
- Levitan, Sar A. and Robert Taggart, "Manpower Programs," The Promise of Greatness, (Cambridge, MA: Harvard University Press, 1976.)
- Louis, Harris and Associates, "A Study of August 1966 Terminations from the Job Corps," in Hearings on the Economic Opportunity Amendments of 1967, U.S. Congress, Senate Committee on Labor and Education, Part I, 90th Congress, First Session, 1967.
- Loury, G. D., "A Dynamic Theory of Racial Income Differences," in Employability Development Teams and Federal Manpower Programs, Richard D. Leone, ed. (Philadelphia: Temple University, 1972.)
- Lamond, A. and Phyllis A. Wallace, eds., Women, Minorities and Employment Discrimination (Lexington, MA: D. C. Heath and Company, 1977.)
- Main, Earl, "A National Evaluation of MDTA Institutional Job Training," Journal of Human Resources, Vol. 3, No. 2., Spring 1968.
- Malkiel, Burton G. and Judith A. Malkiel, "Male-Female Pay Differentials in Professional Employment," American Economic Review, Vol. 63, No. 4, September 1973, pp. 693-705.
- Mangum, Garth L., MDTA: Foundation of Federal Manpower Policy (Baltimore, MD: The John Hopkins Press, 1968.)
- Mangum, Garth L., The Emergence of Manpower Policy (New York: Holt, Rinehart and Winston, Inc., 1969.)
- Mangum, Garth, L. "A Review of Manpower Research," A Review of Industrial Relations Research, Vol. 2 (Madison, WI: Industrial Relations Research Association, 1971.)
- Mangum, Garth L. and Thayne T. Robson, Metropolitan Impact of Manpower Programs: A Four-City Comparison (Salt Lake City, UT: Olympus Publishing Co., 1973.)

Mangum, Garth L. and Robert Taggart, "CETA Service Components: Results and Reasons," National Council on Employment Policy, Washington, DC, June 1981.

Mangum, Garth L. and J. Walsh, A Decade of Manpower Development and Training (Salt Lake City, UT: Olympus Publishing Co., 1973.)

Marshall, Ray, "The Economics of Racial Discrimination: A Survey," Journal of Economic Literature, Vol. 12, No. 3, September 1974, pp. 849-71.

McCall, R. B., Fundamental Statistics for Psychology (New York: Harcourt, Brace, 1975.)

Mincer, Jacob, Schooling, Age and Earnings (New York: Columbia University Press, 1973.)

Mincer, Jacob, Schooling, Experience and Earnings, National Bureau for Economic Research, Washington, DC, 1974.

Mirengoff, William, et.al., The New CETA: Effect of Public Service Programs: Final Report, National Academy of Sciences, Washington, DC, 1980.

Nathan, R. P., et. al., Job Creating Through Public Service Employment: Volume II, Monitoring the Public Service Employment Program, report prepared by the Brookings Institution for the National Commission for Manpower Policy, March 1978.

National Academy of Education, Task Force on Education and Employment, Education for Employment: Knowledge for Action, 1979.

National Commission for Employment Policy, Sixth Annual Report to the Public Employment Program, Vol. 1, prepared for the Office of the Assistant Secretary for Policy, Evaluation and Research, Washington, DC, 1974.

Nunally, Jason C., "The Study of Change in Evaluation Research: Principles Concerning Measurement, Experimental Design, and Analysis," in Handbook of Evaluation Research, Vol. 1, E. L. Struening and M. Guttentag, eds., (Beverly Hills: Sage, 1975.)

Okun, Arthur M., "Conflicting National Goals," in Jobs for Americans, Eli Ginzberg, ed. (Englewood Cliffs, NJ: The American Assembly, 1976), pp. 59-84.

Olympus Research Corporation, "The Total Impact of Manpower Programs: A Four-City Case Study," report to the U. S. Department of Labor, Manpower Administration, Office of Policy Evaluation and Research, Washington, DC, 1971.

O'Neill, David M., The Federal Government and Manpower, American Enterprise Institute, 1973.

Palmer, John L., "Evaluating the Economic Stimulus Package from an Employment and Training Perspective," in Job Creation Through Public Service Employment, Vol. III, Report No. 6, National Commission for Manpower Policy, 1978.

Parnes, Herbert S., "The National Longitudinal Surveys: Lessons for Human Resource Policy," in Current Issues in the Relationship Between Manpower Research and Policy, NCMP, Special Report No. 7, 1976.

Parnes, Herbert S., "The National Longitudinal Surveys: An Interim Assessment" in Manpower Research and Labor Economics, G. Swanson and J. Michaelson, eds. (Beverly Hills: Sage Publications, 1979.)

Parnes, Herbert S., et. al., "The National Longitudinal Surveys: Comprehensive Reports," issued as Manpower Research Monographs Nos. 15, 16, 21 and 24, available from NTIS.

Perry, C. R., et. al., The Impact of Government Manpower Programs in General and on Minorities and Women (Philadelphia: University of Pennsylvania, 1975.)

Pitcher, H. M., "A Sensitivity Analysis to Determine Sample Sizes for Performing Impact Evaluation of the CETA Programs," in Bloch, ed., 1979.

Posavac, E. J. and R. G. Carey, Program Evaluation: Methods and Case Studies (Englewood Cliffs, NJ: Prentice-Hall, 1980.)

Prescott, E. C. and T. F., Cooley, Evaluating the Impact of MDTA Programs on Earnings Under Varying Labor Market Conditions, Final Report, MEL 73-09, University of Pennsylvania, 1972.

Robin, G. D., An Assessment of In-School NYC Projects in Cincinnati and Detroit (Philadelphia: National Analysts, Inc. 1969.)

Rosen, Sherwin, A Theory of Life Earnings, Discussion Paper 72-25, Department of Economics, University of Rochester, Rochester, NY, 1975.

Reid, Clifford, "Some Evidence of the Effect of Manpower Training Programs on the Black/White Wage Differential," Journal of Human Resources, Vol. 11, No. 3, Summer 1976, pp. 402-10.

Schiller, Bradley R., "Welfare: Reforming Our Expectations," The Public Interest, No. 62, 1981.

Sewell, David O., Training the Poor: A Benefit-Cost Analysis of Programs in the United States Antipoverty Program, Industrial Relations Center, Queens University, Kingston, Ontario, 1971.

Siegel, S., Nonparametric Methods for the Behavioral Sciences (New York: McGraw-Hill, 1956.)

Smith, Ralph E., "Analysis of the Efficiency and Equity of Manpower Programs," Ph.D. Dissertation, Georgetown University, Washington, DC, 1971.

Solow, Robert M., "Employment in Inflationary Times," in Employing the Unemployed, Eli Ginzberg, ed. (New York: Basic Books, 1980.), pp. 129-41.

Somers, Gerald G., ed., Retraining the Unemployed (Madison, WI: University of Wisconsin Press, 1968.)

Somers, Gerald, G. and W. D. Wood, eds., Cost-Benefit Analysis of Manpower Policies, Queens University, Kingston, Ontario, 1969.

Stafford, Frank P., "A Decision Theoretic Approach to the Evaluation of Training Programs," in Bloch, ed., 1979.

Stromsdoffer, Ernest W., Review and Synthesis of Cost-Effectiveness Studies of Vocational and Technical Education (Columbus, OH: ERIC Clearing House on Vocational and Technical Education, 1972.)

Stromsdoffer, Ernest W. and Gary Goodfellow, "Cost-Effectiveness Analysis of Four Categorical Employment and Training Programs," unpublished, U.S. Department of Labor Report, 1977.

Stromsdorfer, Ernest W., "The Effectiveness of Youth Programs," in Youth Employment and Public Policy, Bernard E. Anderson and Isabelle V. Sawhill, eds. (Englewood Cliffs, NJ: Prentice-Hall, 1980.)

Suojanen, William, CETA Title I Manpower Programs, Ph.D. Dissertation, Sloan School of Management, MIT, Cambridge, MA, August 1977.

Systems Development Corporation, "Impact of the JOBS Program in Nine Cities," U.S. Department of Labor, Manpower Administration, Office of Program Evaluation and Research, Washington, DC, 1969.

Tella, Dorothy, et. al., A Model for Manpower Training Evaluation (McLean, VA: Planning Research Corporation, 1970.)

Tobin, James, "Stabilization Policy Ten Years After," Brookings Papers on Economic Activity, Vol. 1, The Brookings Institution, Washington, DC, 1980, pp. 23-4.

Ulman, Lloyd, "Manpower Policies and Demand Management," in Jobs for Americans, Eli Ginzberg, ed. (Englewood Cliffs, NJ: Prentice-Hall, 1976), pp. 85-119.

Urban Systems Research and Engineering, Inc., "Impact of Five Rural Concentrated Employment Programs," U.S. Department of Labor, Manpower Administration, Office of Program Evaluation and Research, Washington, DC, 1971.

U.S. Department of Labor, Employment and Training Administration, Work Experience Perspectives: CETA Program Models, U.S. Department of Labor, Washington, DC, 1979

U.S. Department of Labor, Employment and Training Administration, Public Service Employment: CETA Program Models, U.S. Department of Labor, Washington, DC 1978.

U.S. Department of Labor, Employment and Training Administration, Classroom Training -- The OIC Approach: CETA Program Models, U.S. Department of Labor, Washington, DC 1978.

U.S. Department of Labor, Employment and Training Administration, On-The-Job Training: CETA Program Models, U.S. Department of Labor, Washington, DC, 1978.

U.S. Department of Labor, Employment and Training Administration, CETA Program Status and Financial Summary -- Fiscal Year 1980, U.S. Department of Labor, Washington, DC, 1980.

U.S. Department of Labor, Employment and Training Administration, Summary of Participant Characteristics -- Fiscal Year 1980, U.S. Department of Labor, Washington, DC, 1980.

Walther, Regis, et. al., "A Study of the Effectiveness of Selected Out-of-School NYC Programs," Social Research Group, George Washington University, Washington, DC, 1971.

Weisbrod, Barton A., "Benefits of Manpower Programs: Theoretical and Methodological Issues," in Somers and Wood, eds., 1969.

Westat, Inc., The Net Earnings Impact of Public Employment Program (PEP), U.S. Department of Labor, Employment and Training Administration, Office of Program Evaluation, Washington, DC, 1979.

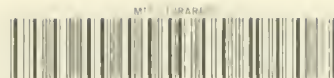
Westat, Inc., Continuous Longitudinal Manpower Survey, Net Impact Report No. 1: Impact on 1977 Earnings of New Fiscal Year 1976 Enrollees in Selected Program Activities, prepared for the U.S. Department of Labor, Employment and Training Administration, Office of Program Evaluation and Research, Washington, DC, March 1981.

Westat, Inc., Continuous Longitudinal Manpower, Follow-Up Report No. 3 (36 Months After Entry) Experiences in the First Two Postprogram Years, With Pre/Post Comparisons for Trainees who Entered CETA during January-June 1975, Vol. 1, descriptive report prepared for the U.S. Department of Labor, Employment and Training Administration, Office of Program Evaluation, Washington, DC, January 1981.

Wirtz, Willard and Harld Goldstein, "Measurement and Analysis of Work Training," Monthly Labor Review, Vol. 98, September 1975, pp. 19-26.

Zoritsky, Jeffrey, et. al., Program Evaluation at the Local Level: A Systems Approach, U.S. Department of Manpower Development, Policy and Evaluation Division, Commonwealth of Massachusetts, 1979.

4953 010



3 9080 004 180 953

BASEMENT
Date Due

Lib-26-67

